

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 10:48:33 ; Search time 21 Seconds
(without alignments)
60.982 Million cell updates/sec

Title: US-09-828-592-10
Perfect score: 58
Sequence: 1 LQEWLKKIKRM 11

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 300 summaries

Database :

Published Applications_AA:*
1: /cgn2-6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
2: /cgn2-6/ptodata/1/pubppaa/PCR_NEW_PUB.pep:*
3: /cgn2-6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2-6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2-6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
6: /cgn2-6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2-6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2-6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2-6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2-6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2-6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2-6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2-6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2-6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	100.0	11	10	US-09-828-592-10
2	41	70.7	11	10	US-09-828-592-11
3	40	69.0	11	10	US-09-828-592-8
4	40	69.0	464	10	US-09-414-834-1
5	40	69.0	934	9	US-09-971-490-2
6	40	69.0	3353	10	US-09-888-615-64
7	39	67.2	823	9	US-10-177-293-486
8	39	67.2	841	10	US-09-895-913A-244
9	39	67.2	841	10	US-09-949-192-27
10	39	67.2	1290	9	US-10-017-128-2
11	37	63.8	114	9	US-09-764-868-1108
12	37	63.8	314	9	US-10-097-065-588
13	37	63.8	429	10	US-09-815-242-10098
14	36	62.1	168	9	US-09-764-868-687
15	36	62.1	424	10	US-09-815-242-10167
16	36	62.1	428	10	US-09-815-242-13851
17	36	62.1	575	9	US-09-971-773-23
18	36	62.1	2441	12	US-10-109-886-10
19	36	62.1	2442	12	US-10-109-886-8

20	35	60.3	11	10	US-09-828-592-7	Sequence 7, App1
21	35	60.3	11	10	US-09-828-592-12	Sequence 12, App1
22	35	60.3	18	9	US-09-865-989-215	Sequence 215, App
23	35	60.3	18	9	US-10-099-574A-215	Sequence 215, App
24	35	60.3	20	9	US-09-741-106-14	Sequence 14, App1
25	35	60.3	51	10	US-09-864-761-47528	Sequence 47528, A
26	35	60.3	59	10	US-09-864-761-48452	Sequence 48452, A
27	35	60.3	89	10	US-09-864-761-46485	Sequence 46485, A
28	35	60.3	379	9	US-10-097-340-139	Sequence 139, App
29	35	60.3	422	9	US-09-374-066A-148	Sequence 148, App
30	35	60.3	422	9	US-10-233-553-25	Sequence 25, App1
31	35	60.3	422	9	US-10-233-553-26	Sequence 26, App1
32	35	60.3	422	9	US-10-233-553-27	Sequence 27, App1
33	35	60.3	422	9	US-10-233-553-28	Sequence 28, App1
34	35	60.3	507	9	US-10-081-872-140	Sequence 140, App
35	35	60.3	1024	9	US-10-211-962-85	Sequence 85, App1
36	35	60.3	1150	10	US-09-946-239-9	Sequence 9, App1
37	35	60.3	1718	9	US-10-223-070-21	Sequence 21, App1
38	35	60.3	2008	9	US-09-736-968A-105	Sequence 105, App
39	35	60.3	2008	10	US-09-978-244A-29	Sequence 29, App1
40	35	60.3	2008	10	US-09-736-969A-12	Sequence 2, App1
41	35	60.3	2008	10	US-09-736-969A-91	Sequence 91, App1
42	35	60.3	2073	9	US-09-978-244A-10	Sequence 88, App1
43	35	60.3	22	9	US-09-865-989-54	Sequence 54, App1
44	34	58.6	22	9	US-10-099-574A-54	Sequence 54, App1
45	34	58.6	22	9	US-09-820-053A-9	Sequence 9, App1
46	34	58.6	23	9	US-09-820-053A-161	Sequence 161, App
47	34	58.6	23	9	US-10-109-171-9	Sequence 9, App1
48	34	58.6	23	9	US-10-109-171-161	Sequence 161, App
49	34	58.6	11	12	US-10-097-079-70	Sequence 12, App1
50	34	58.6	116	9	US-10-177-293-37	Sequence 37, App1
51	34	58.6	123	10	US-10-125-540-408	Sequence 408, App
52	34	58.6	123	10	US-09-754-870A-408	Sequence 408, App
53	34	58.6	135	9	US-09-935-999-95	Sequence 95, App1
54	34	58.6	145	9	US-09-832-355A-115	Sequence 115, App
55	34	58.6	184	9	US-09-967-836-1	Sequence 1, App1
56	34	58.6	195	9	US-10-177-293-387	Sequence 387, App
57	34	58.6	206	9	US-09-864-921-182	Sequence 182, App
58	34	58.6	297	10	US-09-815-242-13468	Sequence 13468, A
59	34	58.6	324	10	US-09-841-132-566	Sequence 566, App
60	34	58.6	359	9	US-09-864-921-99	Sequence 99, App1
61	34	58.6	362	9	US-09-479-040-11	Sequence 11, App1
62	34	58.6	385	9	US-10-205-841-38	Sequence 38, App1
63	34	58.6	445	9	US-10-059-909-14	Sequence 14, App1
64	34	58.6	465	9	US-09-987-021-6	Sequence 6, App1
65	34	58.6	471	9	US-09-900-425A-5	Sequence 5, App1
66	34	58.6	471	9	US-10-079-185-5	Sequence 5, App1
67	34	58.6	485	9	US-10-156-761-9788	Sequence 9788, App
68	34	58.6	503	9	US-10-215-224-4	Sequence 4, App1
69	34	58.6	503	9	US-10-215-224-10	Sequence 10, App1
70	34	58.6	503	9	US-10-225-060-18	Sequence 18, App1
71	34	58.6	503	9	US-09-213-888-7	Sequence 7, App1
72	34	58.6	540	9	US-09-328-877A-10	Sequence 10, App1
73	34	58.6	540	9	US-09-328-877A-6	Sequence 6, App1
74	34	58.6	540	9	US-09-328-877A-5	Sequence 5, App1
75	34	58.6	540	9	US-09-328-877A-4	Sequence 4, App1
76	34	58.6	545	9	US-09-213-888-6	Sequence 6, App1
77	34	58.6	545	9	US-09-328-877A-6	Sequence 6, App1
78	34	58.6	553	9	US-09-213-888-5	Sequence 5, App1
79	34	58.6	553	9	US-09-328-877A-5	Sequence 5, App1
80	34	58.6	559	9	US-09-213-888-9	Sequence 9, App1
81	34	58.6	559	9	US-09-328-877A-9	Sequence 9, App1
82	34	58.6	589	9	US-09-213-888-8	Sequence 8, App1
83	34	58.6	589	9	US-09-328-877A-8	Sequence 8, App1
84	34	58.6	592	9	US-09-213-888-4	Sequence 4, App1
85	34	58.6	626	9	US-09-213-888-21	Sequence 21, App1
86	34	58.6	626	9	US-09-328-877A-3	Sequence 3, App1
87	34	58.6	627	9	US-09-328-877A-3	Sequence 3, App1
88	34	58.6	666	9	US-09-213-888-27	Sequence 27, App1
89	34	58.6	666	9	US-09-328-877A-27	Sequence 27, App1
90	34	58.6	666	9	US-09-328-877A-27	Sequence 27, App1
91	34	58.6	666	9	US-09-328-877A-27	Sequence 27, App1
92	34	58.6	666	9	US-09-328-877A-27	Sequence 27, App1

93	58.6	669	9	US-09-213-888-25	Sequence 25, Appl	166	33	56.9	747	9	US-10-123-902-212	Sequence 212, App
94	58.6	669	9	US-09-328-877A-25	Sequence 25, Appl	167	33	56.9	747	9	US-10-123-908-212	Sequence 212, App
95	58.6	748	9	US-09-291-417-24	Sequence 24, Appl	168	33	56.9	747	9	US-10-123-909-212	Sequence 212, App
96	58.6	786	9	US-09-291-417-23	Sequence 23, Appl	169	33	56.9	747	9	US-10-123-910-212	Sequence 212, App
97	58.6	934	10	US-09-841-739-11	Sequence 11, Appl	170	33	56.9	747	9	US-10-124-813-212	Sequence 212, App
98	58.6	934	9	US-09-912-697-10	Sequence 10, Appl	171	33	56.9	747	9	US-10-124-817-212	Sequence 212, App
99	58.6	934	9	US-09-760-285-20	Sequence 20, Appl	172	33	56.9	747	9	US-10-124-824-212	Sequence 212, App
100	58.6	934	10	US-09-788-657-19	Sequence 19, Appl	173	33	56.9	747	9	US-10-125-922-212	Sequence 212, App
101	58.6	1001	9	US-09-291-417-31	Sequence 31, Appl	174	33	56.9	747	9	US-10-125-924-212	Sequence 212, App
102	58.6	1005	9	US-10-177-293-254	Sequence 254, App	175	33	56.9	747	9	US-10-127-825A-212	Sequence 212, App
103	58.6	1024	9	US-09-864-921-97	Sequence 97, Appl	176	33	56.9	747	9	US-10-127-829A-212	Sequence 212, App
104	58.6	1024	9	US-10-156-733-2	Sequence 2, Appl	177	33	56.9	747	9	US-10-127-835A-212	Sequence 212, App
105	58.6	1024	10	US-09-841-739-2	Sequence 5, Appl	178	33	56.9	747	9	US-10-127-839A-212	Sequence 212, App
106	58.6	1204	10	US-09-841-739-5	Sequence 17, Appl	179	33	56.9	747	9	US-10-127-901A-212	Sequence 212, App
107	58.6	1263	10	US-09-864-761-6846	Sequence 17, Appl	180	33	56.9	747	9	US-10-128-693A-212	Sequence 212, App
108	58.6	1494	9	US-10-195-144-17	Sequence 46862, A	181	33	56.9	747	9	US-10-131-813A-212	Sequence 212, App
109	58.6	1706	10	US-09-864-761-46862	Sequence 46862, A	182	33	56.9	747	9	US-10-131-818A-212	Sequence 212, App
110	58.6	4563	9	US-09-870-759-128	Sequence 128, App	183	33	56.9	747	9	US-10-131-822A-212	Sequence 212, App
111	58.6	4563	9	US-09-802-640-32	Sequence 32, Appl	184	33	56.9	747	9	US-10-131-824A-212	Sequence 212, App
112	58.6	18	9	US-09-865-989-205	Sequence 205, App	185	33	56.9	747	9	US-10-131-830A-212	Sequence 212, App
113	56.9	18	9	US-10-099-574A-205	Sequence 205, App	186	33	56.9	747	9	US-10-131-837A-212	Sequence 212, App
114	56.9	83	9	US-10-091-572-237	Sequence 237, App	187	33	56.9	747	9	US-10-137-872A-212	Sequence 212, App
115	56.9	83	9	US-09-764-891-3348	Sequence 3348, Ap	188	33	56.9	747	9	US-10-140-860-212	Sequence 212, App
116	56.9	107	9	US-10-169-048-46	Sequence 46, Appl	189	33	56.9	747	9	US-10-142-417-212	Sequence 212, App
117	56.9	160	9	US-09-738-626-4352	Sequence 4352, Ap	190	33	56.9	747	9	US-10-147-500-212	Sequence 212, App
118	56.9	226	9	US-09-738-626-3601	Sequence 3601, Ap	191	33	56.9	747	9	US-10-147-502-212	Sequence 212, App
119	56.9	246	10	US-09-925-302-534	Sequence 534, App	192	33	56.9	747	9	US-10-147-515-212	Sequence 212, App
120	56.9	315	9	US-10-106-698-5557	Sequence 5557, Ap	193	33	56.9	747	9	US-10-147-517-212	Sequence 212, App
121	56.9	393	9	US-10-128-714-3421	Sequence 3421, Ap	194	33	56.9	747	9	US-10-147-518-212	Sequence 212, App
122	56.9	398	9	US-10-106-698-5537	Sequence 5537, Ap	195	33	56.9	747	9	US-10-147-526-212	Sequence 212, App
123	56.9	409	10	US-09-925-301-1089	Sequence 1089, Ap	196	33	56.9	747	9	US-10-147-527-212	Sequence 212, App
124	56.9	433	9	US-10-255-969-6	Sequence 6, Appl	197	33	56.9	747	9	US-10-152-395-212	Sequence 212, App
125	56.9	433	12	US-10-139-262-6	Sequence 8421, Ap	198	33	56.9	747	9	US-10-157-782A-212	Sequence 212, App
126	56.9	451	9	US-10-128-714-8421	Sequence 8421, Ap	199	33	56.9	747	9	US-10-160-840-212	Sequence 212, App
127	56.9	484	9	US-10-114-810-2	Sequence 2, Appl	200	33	56.9	747	9	US-10-121-056-212	Sequence 212, App
128	56.9	549	9	US-09-883-152-23	Sequence 23, Appl	201	33	56.9	747	9	US-10-121-061-212	Sequence 212, App
129	56.9	533	9	US-09-738-626-3655	Sequence 3655, Ap	202	33	56.9	747	9	US-10-123-235-212	Sequence 212, App
130	56.9	611	9	US-09-893-519A-28	Sequence 28, Appl	203	33	56.9	747	9	US-10-124-818-212	Sequence 212, App
131	56.9	674	10	US-09-765-272-200	Sequence 200, App	204	33	56.9	747	9	US-10-125-926A-212	Sequence 212, App
132	56.9	747	9	US-10-028-072-212	Sequence 212, App	205	33	56.9	747	9	US-10-125-930A-212	Sequence 212, App
133	56.9	747	9	US-10-121-049-212	Sequence 212, App	206	33	56.9	747	9	US-10-127-831A-212	Sequence 212, App
134	56.9	747	9	US-10-123-904-212	Sequence 212, App	207	33	56.9	747	9	US-10-127-837A-212	Sequence 212, App
135	56.9	747	9	US-10-140-470-212	Sequence 212, App	208	33	56.9	747	9	US-10-127-838B-212	Sequence 212, App
136	56.9	747	9	US-10-175-746-212	Sequence 212, App	209	33	56.9	747	9	US-10-127-842A-212	Sequence 212, App
137	56.9	747	9	US-10-176-918-212	Sequence 212, App	210	33	56.9	747	9	US-10-127-843A-212	Sequence 212, App
138	56.9	747	9	US-10-176-921-212	Sequence 212, App	211	33	56.9	747	9	US-10-127-845A-212	Sequence 212, App
139	56.9	747	9	US-10-137-865-212	Sequence 212, App	212	33	56.9	747	9	US-10-127-846A-212	Sequence 212, App
140	56.9	747	9	US-10-140-474-212	Sequence 212, App	213	33	56.9	747	9	US-10-127-848A-212	Sequence 212, App
141	56.9	747	9	US-10-142-431-212	Sequence 212, App	214	33	56.9	747	9	US-10-127-849A-212	Sequence 212, App
142	56.9	747	9	US-10-143-114-212	Sequence 212, App	215	33	56.9	747	9	US-10-127-850A-212	Sequence 212, App
143	56.9	747	9	US-10-140-002-212	Sequence 212, App	216	33	56.9	747	9	US-10-127-851A-212	Sequence 212, App
144	56.9	747	9	US-10-142-419-212	Sequence 212, App	217	33	56.9	747	9	US-10-128-684A-212	Sequence 212, App
145	56.9	747	9	US-10-123-262-212	Sequence 212, App	218	33	56.9	747	9	US-10-128-686A-212	Sequence 212, App
146	56.9	747	9	US-10-142-423-212	Sequence 212, App	219	33	56.9	747	9	US-10-128-690A-212	Sequence 212, App
147	56.9	747	9	US-10-121-050-212	Sequence 212, App	220	33	56.9	747	9	US-10-128-691A-212	Sequence 212, App
148	56.9	747	9	US-10-141-755-212	Sequence 212, App	221	33	56.9	747	9	US-10-128-691A-212	Sequence 212, App
149	56.9	747	9	US-10-143-032-212	Sequence 212, App	222	33	56.9	747	9	US-10-131-819A-212	Sequence 212, App
150	56.9	747	9	US-10-123-108-212	Sequence 212, App	223	33	56.9	747	9	US-10-131-829A-212	Sequence 212, App
151	56.9	747	9	US-10-123-236-212	Sequence 212, App	224	33	56.9	747	9	US-10-131-836A-212	Sequence 212, App
152	56.9	747	9	US-10-123-261-212	Sequence 212, App	225	33	56.9	747	9	US-10-137-868-212	Sequence 212, App
153	56.9	747	9	US-10-140-921-212	Sequence 212, App	226	33	56.9	747	9	US-10-146-729-212	Sequence 212, App
154	56.9	747	9	US-10-140-928-212	Sequence 212, App	227	33	56.9	747	9	US-10-146-791-212	Sequence 212, App
155	56.9	747	9	US-10-121-045-212	Sequence 212, App	228	33	56.9	747	9	US-10-147-484-212	Sequence 212, App
156	56.9	747	9	US-10-123-292-212	Sequence 212, App	229	33	56.9	747	9	US-10-147-492-212	Sequence 212, App
157	56.9	747	9	US-10-123-903-212	Sequence 212, App	230	33	56.9	747	9	US-10-147-508-212	Sequence 212, App
158	56.9	747	9	US-10-124-819-212	Sequence 212, App	231	33	56.9	747	9	US-10-147-512-212	Sequence 212, App
159	56.9	747	9	US-10-124-822-212	Sequence 212, App	232	33	56.9	747	9	US-10-158-782-212	Sequence 212, App
160	56.9	747	9	US-10-124-825-212	Sequence 212, App	233	33	56.9	747	9	US-10-175-735-212	Sequence 212, App
161	56.9	747	9	US-10-160-498-212	Sequence 212, App	234	33	56.9	747	9	US-10-123-905-212	Sequence 212, App
162	56.9	747	9	US-10-121-041-212	Sequence 212, App	235	33	56.9	747	9	US-10-123-907-212	Sequence 212, App
163	56.9	747	9	US-10-121-043-212	Sequence 212, App	236	33	56.9	747	9	US-10-124-815-212	Sequence 212, App
164	56.9	747	9	US-10-121-047-212	Sequence 212, App	237	33	56.9	747	9	US-10-125-921A-212	Sequence 212, App
165	56.9	747	9	US-10-123-215-212	Sequence 212, App	238	33	56.9	747	9	US-10-125-926A-212	Sequence 212, App


```

; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-592-8

```

Query Match	69.0%	Score 40;	DB 10;	Length 11;
Best Local Similarity	63.6%;	Pred. No. 5.5;		
Matches 7;	Conservative 3;	Mismatches 1;	Indels 0;	Gaps 0;

Qy	1 LQEWLKKLKM 11
	:
Db	1 LQEWLDELEEM 11

```

: RESULT 4
: US-09-414-834-1
: Sequence 1, Application US/09414834
: Patent No. US20020076413A1
: GENERAL INFORMATION:
: APPLICANT: O'Reilly, Michael S.
: APPLICANT: Folkman, M. Judah
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING ANGIOGENESIS
: FILE REFERENCE: 07555.0001
: CURRENT APPLICATION NUMBER: US/09/414,834
: NUMBER OF SEQ ID NOS: 1
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 464
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Antithrombin III
: US-09-414-834-1

```

	Query Match	Score	DB	Length
	Best Local Similarity	69.0%	40	464
		63.6%	Pred. No. 1.9e+02	
Matches	7, Conservative	3, Mismatches	1, Indels	0, Gaps
QY	1 LQEWLKLKKM 11			
	: :			
Db	336 LQEWLDELLEM 346			

QY	1	LQEWLKKLKM	11
		: ::	
Db	336	LQEWLDELEEM	346

```

RESULT 5
US-09-971-490-2
Sequence 2, Application US/09971490
Publication No. US20030032091A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
APPLICANT: Curtis, Rory A. J.
APPLICANT: Bandaru, Rajasekhar
TITLE OF INVENTION: 48120, 23479 AND 46689, NOVEL HUMAN HYDROLASES
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: 10448-100001
CURRENT APPLICATION NUMBER: US/09/971,490
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 60/238,170
PRIOR FILING DATE: 2000-10-05
PRIOR APPLICATION NUMBER: 60/237,991
PRIOR FILING DATE: 2000-10-05
NUMBER OF SEQ ID NOS: 20
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 934
TYPE: PRT
ORGANISM: Homo sapiens
US-09-971-490-2

```

Query Match	69.0%	Score 40:	DB 9:	Length 934:
Best Local Similarity	87.5%	Pred. No.	3.8e+02:	
Matches 7; Conservative	1;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	LQEWLKKL	8
		11:11111	
Db	55	LQDWLKKL	62

```

US-09-888-615-64
RESULT 6
Sequence 64, Application US/09888615
Patent No. US20020064856A1
GENERAL INFORMATION:
APPLICANT: PLOWMAN, GREGORY
APPLICANT: WHYTE, DAVID
APPLICANT: CAENEPEEL, SEAN
APPLICANT: CHARVETZAK, GLEN
APPLICANT: MANNING, GERARD
APPLICANT: SUDRANSAM, SUCHA
TITLE OF INVENTION: NOVEL PROTEASES
FILE REFERENCE: 038602/1214
CURRENT APPLICATION NUMBER: US/09/888,615
CURRENT FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 60/214,047
PRIOR FILING DATE: 2000-06-26
NUMBER OF SEQ ID NOS: 150
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 64
LENGTH: 3353
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (1891)
OTHER INFORMATION: Any amino acid
US-09-888-615-64

```

Query Match	69.0%	Score 40:	DB 10:	Length 3353:
Best Local Similarity	87.5%	Pred. No. 1.3e+03:		
Matches	7:	Conservative	1:	Mismatches 0:
				Indels 0:
				Gaps 0:
QY	1	LDQWIKKL	8	
		:		
Db	1460	LDQWIKKL	1467	

```
QY      1 LQEWLKKL 8
         ||:|||||
Db      1460 LQDWLKKL 1467
```

```

1      RESULT 7
2      US-10-177-293-486
3      ; Sequence 486, Application US/10177293
4      ; Publication No. US20030124128A1
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Lillie, James
7      ; APPLICANT: Glatt, Karen
8      ; APPLICANT: Zhao, Xumei
9      ; APPLICANT: Gannavarpu, Manjula
10     ; APPLICANT: Kamatkar, Shubhangi
11     ; APPLICANT: Mertens, Maureen
12     ; APPLICANT: Myer, Vic
13     ; APPLICANT: Wang, Youzhen
14     ; APPLICANT: Xu, Yongyao
15     ; APPLICANT: Hoersch, Sebastian
16     ; APPLICANT: Monahan, John
17     ; APPLICANT: Meyers, Rachel E.
18     ; APPLICANT: Bast Jr., Robert C.
19     ; APPLICANT: Hortobagyi, Gabriel N.
20     ; APPLICANT: Pusztai, Lajos
21     ; APPLICANT: Meric, Funda
22     ; APPLICANT: Sahin, Aysegul
23     ; APPLICANT: Mills, Gordon B.
24     ; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT
25     ; FILE REFERENCE: MRI-038
26     ; CURRENT APPLICATION NUMBER: US/10/177,293
27     ; PRIOR FILING DATE: 2002-06-21
28     ; PRIOR APPLICATION NUMBER: US 60/299,887

```



```
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 486
; LENGTH: 798
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-486
```

```
Query Match      67.2%; Score 39; DB 9; Length 798;
Best Local Similarity 60.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      1 LQEWLKKLKK 10
        |||:||||
Db      485 LQWIKLKK 494
```

```
RESULT 8
US-09-895-913A-244
; Sequence 244, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020160456A1el Helicobacter Polypeptides in the
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 244
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-895-913A-244
```

```
Query Match      67.2%; Score 39; DB 9; Length 823;
Best Local Similarity 60.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      1 LQEWLKKLKK 10
        |||:||||
Db      246 MQEWIKDLKE 255
```

```
RESULT 9
US-09-949-192-27
; Sequence 27, Application US/09949192
; Patent No. US20020142292A1
; GENERAL INFORMATION:
; APPLICANT: Parham, Christl L.
; APPLICANT: Gorman, Daniel L.
; APPLICANT: Kurata, Hirokazu
; APPLICANT: Arai, Naoko
; APPLICANT: Sana, Theodore R.
```

```
; APPLICANT: Mattson, Jeanine D.
; APPLICANT: Murphy, Erin E.
; APPLICANT: Savkoor, Chetan
; APPLICANT: Grein, Jeffery
; APPLICANT: Smith, Kathleen M.
; APPLICANT: McElanahan, Terrill K.
; TITLE OF INVENTION: MAMMALIAN GENES: RELATED REAGENTS AND METHODS
; FILE REFERENCE: DX01169K
; CURRENT APPLICATION NUMBER: US/09/949,192
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/231,267
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-949-192-27
```

```
Query Match      67.2%; Score 39; DB 10; Length 841;
Best Local Similarity 60.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      1 LQEWLKKLKK 10
        |||:||||
Db      528 LQWIKLKK 537
```

```
RESULT 10
US-10-017-128-2
; Sequence 2, Application US/10017128
; Publication No. US20030124536A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT FOR VASCULAR DISEASE
; FILE REFERENCE: WMT-001
; CURRENT APPLICATION NUMBER: US/10/017,128
; CURRENT FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/306,941
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/315,572
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/327,488
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1290
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-128-2
```

```
Query Match      67.2%; Score 39; DB 9; Length 1290;
Best Local Similarity 45.5%; Pred. No. 7.2e+02;
Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      1 LQEWLKKLKK 11.
        |||:||||:
Db      919 LQDWVKIKREV 929
```

```
RESULT 11
US-09-764-868-1108
; Sequence 1108, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT732
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
```

```
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1108
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (26)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (36)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-1108

Query Match
Best Local Similarity 63.8%; Score 37; DB 9; Length 114;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 4 WLKRLKKM 11
Db 2 WLKRLKKM 9

RESULT 12
US-10-097-065-588
; Sequence 588, Application US/10097065
; Publication No. US20030055236A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A. et al.
; TITLE OF INVENTION: 110 Human Secreted Proteins
; FILE REFERENCE: P2021P1
; CURRENT APPLICATION NUMBER: US/10/097,065
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: PCT/US98/27059
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,007
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,057
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,006
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,369
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,367
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,169
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,053
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,064
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,054
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,008
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,365
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 588
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (88)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
```

```
; LOCATION: (94)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (102)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (118)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (121)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (236)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-097-065-588

Query Match
Best Local Similarity 63.8%; Score 37; DB 9; Length 314;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 2 QEWLKLKK 10
Db 37 QSMKMKLKK 45

RESULT 13
US-09-815-242-10098
; Sequence 10098, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10098
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10098

Query Match
Best Local Similarity 63.8%; Score 37; DB 10; Length 429;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 3 EWLKLKKM 11
```

Db 226 EMLKIRKI 234

|||||:

RESULT 14

US-09-764-868-687
; Sequence 687, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 687
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-868-687

Query Match 62.1%; Score 36; DB 9; Length 168;
Best Local Similarity 45.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 LQEWLKKRK 11
Db 53 VKMWKLRKM 63

RESULT 15

US-09-815-242-10167
; Sequence 10167, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10167
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10167

Query Match 62.1%; Score 36; DB 10; Length 424;

Best Local Similarity 50.0%; Pred. No. 6.9e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 LQEWLKKRK 10
Db 97 VQEWVWKIRK 106

RESULT 16

US-09-815-242-13851
; Sequence 13851, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13851
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Salmoneella typhi
US-09-815-242-13851

Query Match 62.1%; Score 36; DB 10; Length 428;
Best Local Similarity 50.0%; Pred. No. 7e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 LQEWLKKRK 10
Db 101 VQEWVWKIRK 110

RESULT 17

US-09-971-773-23
; Sequence 23, Application US/0971773
; Publication No. US20030115614A1
; GENERAL INFORMATION:
; APPLICANT: Yutaka KANDA
; APPLICANT: Mitsuo SATOH
; APPLICANT: Kazuyasu NAKAMURA
; APPLICANT: Kazuhisa UCHIDA
; APPLICANT: Toyohide SHINKAWA
; APPLICANT: Naoko YAMANE
; APPLICANT: Motoo YAMASAKI
; APPLICANT: NO. US20030115614A1uo HANAI
; TITLE OF INVENTION: ANTIBODY COMPOSITION-PRODUCING CELL
; FILE REFERENCE: 249-202
; CURRENT APPLICATION NUMBER: US/09/971,773

Fri Jul 11 16:11:41 2003

us-09-828-592-10.rapb

CURRENT FILING DATE: 2002-08-30
 PRIOR APPLICATION NUMBER: JP 2000-308526
 PRIOR FILING DATE: 2000-10-06
 PRIOR APPLICATION NUMBER: US 60/268,926
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 73
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 23
 LENGTH: 575
 TYPE: PRT
 ORGANISM: Citicellus griseus
 US-09-971-773-23

Query Match 62.1%; Score 36; DB 9; Length 575;
 Best Local Similarity 72.7%; Pred. No. 9.2e+02; Indels 0; Gaps 0;
 Matches 8; Conservative 1; Mismatches 2;

OY 1 LOEWLKKLKK 11
 11 111111
 DB 130 LOSELKKLKKL 140

RESULT 18
 US-10-109-886-8 Application US/10109886
 Sequence 8, Application US/10109886
 Patent No. US20020119499A1
 GENERAL INFORMATION:
 APPLICANT: TANABE SEIYAKU CO. LTD.
 APPLICANT: TANIGUCHI, Junko
 APPLICANT: MIZUKAMI, Junko
 TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND
 TITLE OF INVENTION: ANTAGONIST TO PPAR
 FILE REFERENCE: TANIGUCHI-6
 CURRENT FILING DATE: 2002-04-01
 PRIOR APPLICATION NUMBER: 09/514,247
 PRIOR FILING DATE: 2000-02-28
 PRIOR APPLICATION NUMBER: PCT/JP98/03734
 PRIOR FILING DATE: 1998-08-24
 PRIOR APPLICATION NUMBER: JP231084/1997
 PRIOR FILING DATE: 1997-08-27
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 8
 LENGTH: 2441
 TYPE: PRT
 ORGANISM: mouse
 US-10-109-886-8

Query Match 62.1%; Score 36; DB 12; Length 2441;
 Best Local Similarity 75.0%; Pred. No. 3.7e+03; Indels 0; Gaps 0;
 Matches 6; Conservative 1; Mismatches 1;

OY 1 LOEWLKKL 8
 1111 11
 DB 1500 LOEWYKKM 1507

RESULT 19
 US-10-109-886-10 Application US/10109886
 Sequence 10, Application US/10109886
 Patent No. US20020119499A1
 GENERAL INFORMATION:
 APPLICANT: TANABE SEIYAKU CO. LTD.
 APPLICANT: TANIGUCHI, Junko
 APPLICANT: MIZUKAMI, Junko
 TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND
 TITLE OF INVENTION: ANTAGONIST TO PPAR
 FILE REFERENCE: TANIGUCHI-6
 CURRENT FILING DATE: 2002-04-01
 PRIOR APPLICATION NUMBER: 09/514,247
 PRIOR FILING DATE: 2000-02-28

PRIOR APPLICATION NUMBER: PCT/JP98/03734
 PRIOR FILING DATE: 1998-08-24
 PRIOR APPLICATION NUMBER: JP231084/1997
 PRIOR FILING DATE: 1997-08-27
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 10
 LENGTH: 2442
 TYPE: PRT
 ORGANISM: human
 US-10-109-886-10

Query Match 62.1%; Score 36; DB 12; Length 2442;
 Best Local Similarity 75.0%; Pred. No. 3.7e+03; Indels 0; Gaps 0;
 Matches 6; Conservative 1; Mismatches 1;

OY 1 LOEWLKKL 8
 1111 11
 DB 1499 LOEWYKKM 1506

RESULT 20
 US-09-828-592-7 Application US/09828592
 Sequence 7, Application US/09828592
 Patent No. US20010055591A1
 GENERAL INFORMATION:
 APPLICANT: Walston, Timothy
 APPLICANT: Cooper, Scott
 APPLICANT: Revzale, Alireza
 TITLE OF INVENTION: ANTIHROMBIN H-HELIX MUTANTS
 FILE REFERENCE: 7869.10US01
 CURRENT FILING DATE: 2001-04-06
 PRIOR APPLICATION NUMBER: 60/195,872
 PRIOR FILING DATE: 2000-04-07
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 7
 LENGTH: 11
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-828-592-7

Query Match 60.3%; Score 35; DB 10; Length 11;
 Best Local Similarity 60.0%; Pred. No. 30; Indels 0; Gaps 0;
 Matches 6; Conservative 2; Mismatches 2;

OY 1 LOEWLKKL 10
 1111 11
 DB 1 LKKWLKKFK 10

RESULT 21
 US-09-828-592-12 Application US/09828592
 Sequence 12, Application US/09828592
 Patent No. US20010055591A1
 GENERAL INFORMATION:
 APPLICANT: Walston, Timothy
 APPLICANT: Cooper, Scott
 APPLICANT: Revzale, Alireza
 TITLE OF INVENTION: ANTIHROMBIN H-HELIX MUTANTS
 FILE REFERENCE: 7869.10US01
 CURRENT FILING DATE: 2001-04-06
 PRIOR APPLICATION NUMBER: 60/195,872
 PRIOR FILING DATE: 2000-04-07
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 12
 LENGTH: 11
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-828-592-12

Query Match 60.3%; Score 35; DB 10; Length 11;
Best Local Similarity 60.0%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LOEWLKLKK 10
DB 1 LOEWLELEEE 10

RESULT 22

US-09-865-989-215
; Sequence 215, Application US/09865989
; Publication No. US20030008827A1
; GENERAL INFORMATION:

APPLICANT: Dasseux, Jean-Louis

Sekul, Renate

Buttner, Klaus

Cornut, Isabelle

Metz, Gunther

TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS

NUMBER OF SEQUENCES: 258

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036-2811

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/865,989

FILING DATE: 25-MAY-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/465,719

FILING DATE: 17-DEC-1999

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 009196-0006-999

TELEPHONE: 650-493-4935

TELEFAX: 650-493-5556

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 215:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: NO. US20030008827A1e

FEATURE:

NAME/KEY: Other

LOCATION: 1..18

OTHER INFORMATION: N-terminal acetylated and

C-terminal amidated

SEQUENCE DESCRIPTION: SEQ ID NO: 215:

US-09-865-989-215

Query Match 60.3%; Score 35; DB 9; Length 18;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LOEWLKLKK 9
DB 10 LEEWKOKLK 18

RESULT 23

US-10-099-574A-215
; Sequence 215, Application US/10099574A
; Publication No. US20030060604A1
; GENERAL INFORMATION:

APPLICANT: Dasseux, Jean-Louis

Sekul, Renate

Buttner, Klaus

Cornut, Isabelle

Metz, Gunther

TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS

NUMBER OF SEQUENCES: 254

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036-2811

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/099,574A

FILING DATE: 29-SEP-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 009196-0005-999

TELEPHONE: 650-493-4935

TELEFAX: 650-493-5556

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 215:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: NO. US20030060604A1e

FEATURE:

NAME/KEY: Other

LOCATION: 1..18

OTHER INFORMATION: N-terminal acetylated and

C-terminal amidated

US-10-099-574A-215

Query Match 60.3%; Score 35; DB 9; Length 18;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LOEWLKLKK 9
DB 10 LEEWKOKLK 18

RESULT 24

US-09-741-106-14
; Sequence 14, Application US/09741106
; Publication No. US20020197667A1
; GENERAL INFORMATION:

APPLICANT: Innis, Michael

Creasey, Abila

TITLE OF INVENTION: Chimeric Proteins

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/741,106
FILING DATE: 12-Dec-2000
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/286,521
FILING DATE: 1994-08-05
ATTORNEY/AGENT INFORMATION:
NAME: Saveriede, Paul B.
REGISTRATION NUMBER: 36,914
REFERENCE/DOCKET NUMBER: 0990,001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2585
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-741-106-14
Query Match 60.3%; Score 35; DB 9; Length 20;
Best Local Similarity 60.0%; Pred. No. 52;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
DB 1 OEWLKRLKK 10
5 LKWLKMKFK 14
RESULT 25
US-09-864-761-47528
Sequence 47528, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 47528
LENGTH: 51
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC022205.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
OTHER INFORMATION: EST_HUMAN HIT: AI377428.1, EVALUATE 3.00e-05
US-09-864-761-47528
Query Match 60.3%; Score 35; DB 10; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 2 OEWLK 7
33 OEWLK 38
RESULT 26
US-09-864-761-48452
Sequence 48452, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aecmica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 48452
LENGTH: 59
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC007021.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
OTHER INFORMATION: EST_HUMAN HIT: W23667.1, EVALUE 4.00e-10
OTHER INFORMATION: SWISSPROT HIT: P24720, EVALUE 5.10e-01
US-09-864-761-48452

Query Match 60.3%; Score 35; DB 10; Length 59;
Best Local Similarity 54.5%; Pred. NO. 1.5e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Oy 1 LQEWLKKKK 11
Db 29 MEMLITLTKI 39

RESULT 27
US-09-864-761-46485
Sequence 46485, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aesomica-X-1
CURRENT FILING DATE: US/09/864,761
PRIOR APPLICATION NUMBER: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 46485
LENGTH: 89
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AI049839.2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2
OTHER INFORMATION: SWISSPROT HIT: P05154, EVALUE 2.00e-48
OTHER INFORMATION: EST_HUMAN HIT: BE937835.1, EVALUE 2.00e-47
US-09-864-761-46485

Query Match 60.3%; Score 35; DB 10; Length 89;
Best Local Similarity 60.0%; Pred. NO. 2.2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Oy 1 LQEWLKKKK 10
Db 80 LRKWLKMKK 89

RESULT 28
US-10-097-340-139
Sequence 139, Application US/10097340
Publication No. US20030087250A1
GENERAL INFORMATION:
APPLICANT: John MONAHAN
APPLICANT: Manjula GANNAVARAPU
APPLICANT: Sebastian HOERSCH
APPLICANT: Shubhangi KAMATKAR
APPLICANT: Steve G. KOVATS
APPLICANT: Rachel E. MEYERS
APPLICANT: Michael MORRISSEY
APPLICANT: Peter OLANDT
APPLICANT: Ami SEN
APPLICANT: Peter VEIBY
APPLICANT: Gordon B. MILLS
APPLICANT: Robert C. BAST, JR.
APPLICANT: Karen LU
APPLICANT: Rosemarie SCHMANDT
APPLICANT: Xumei ZHAO
APPLICANT: Karen GLATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins for The Identification,
FILE REFERENCE: MRI-030
CURRENT FILING DATE: US/10/097,340
PRIOR APPLICATION NUMBER: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-26

us-09-828-592-10.rapb

Fri Jul 11 16:11:41 2003

PRIOR APPLICATION NUMBER: 60/276,026
 PRIOR FILING DATE: 2001-03-14
 PRIOR APPLICATION NUMBER: 60/324,967
 PRIOR FILING DATE: 2001/09/26
 PRIOR APPLICATION NUMBER: 60/311,732
 PRIOR FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: 60/325,102
 PRIOR FILING DATE: 2001-09-26
 PRIOR APPLICATION NUMBER: 60/323,580
 PRIOR FILING DATE: 2001-09-19
 NUMBER OF SEQ ID NOS: 363
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 139
 LENGTH: 379
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-097-340-139

Query Match 60.3%; Score 35; DB 9; Length 379;
 Best Local Similarity 60.0%; Pred. No. 8.7e+02; Indels 0; Gaps 0;
 Matches 6; Conservative 2; Mismatches 2;

OY 1 LOEWLKKLK 10
 DB 260 LKKLKKKK 269

RESULT 29
 US-09-374-046A-148
 Sequence 148, Application US/09374046A
 Publication No. US20030096951A1

GENERAL INFORMATION:
 APPLICANT: Jacobs, Kenneth
 APPLICANT: McCoy, John M.
 APPLICANT: Lavallee, Edward R.
 APPLICANT: Collins-Racie, Lisa A.
 APPLICANT: Evans, Cheryl
 APPLICANT: Merberg, David
 APPLICANT: Treacy, Maurice
 APPLICANT: Agostino, Michael J.
 APPLICANT: Steindinger II, Robert J.
 APPLICANT: Spaulding, Vikki
 APPLICANT: Wong, Gordon G.
 APPLICANT: Clark, Hilary
 APPLICANT: Fechtel, Kim
 APPLICANT: Genetics Institute, Inc.
 TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 FILE REFERENCE: GI 6075-83A
 CURRENT FILING DATE: 1999-08-13
 NUMBER OF SEQ ID NOS: 240
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 148
 LENGTH: 422
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-374-046A-148

Query Match 60.3%; Score 35; DB 9; Length 422;
 Best Local Similarity 77.8%; Pred. No. 9.6e+02; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 2;

OY 1 LOEWLKKLK 9
 DB 73 LOEWLKKK 81

RESULT 30
 US-10-233-553-25
 Sequence 25, Application US/10233553
 Publication No. US20030125285A1
 GENERAL INFORMATION:
 APPLICANT: NIPPON SHINYAKU CO., LTD.

APPLICANT: HIRABAYASHI, Kazuko
 APPLICANT: YANO, Junichi

TITLE OF INVENTION: Method of examining the efficacy of therapy with nucleic acid
 FILE REFERENCE: B-345
 CURRENT FILING DATE: 2002-12-19
 PRIOR APPLICATION NUMBER: JP 2001-267385
 PRIOR FILING DATE: 2001-09-04
 NUMBER OF SEQ ID NOS: 31
 SEQ ID NO: 25
 LENGTH: 422
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-233-553-25

Query Match 60.3%; Score 35; DB 9; Length 422;
 Best Local Similarity 77.8%; Pred. No. 9.6e+02; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 2;

OY 1 LOEWLKKLK 9
 DB 73 LOEWLKKK 81

RESULT 31
 US-10-233-553-26
 Sequence 26, Application US/10233553
 Publication No. US20030125285A1

GENERAL INFORMATION:
 APPLICANT: NIPPON SHINYAKU CO., LTD.
 APPLICANT: HIRABAYASHI, Kazuko
 APPLICANT: YANO, Junichi
 TITLE OF INVENTION: Method of examining the efficacy of therapy with nucleic acid
 FILE REFERENCE: B-345
 CURRENT FILING DATE: 2002-12-19
 CURRENT APPLICATION NUMBER: JP 2001-267385
 PRIOR FILING DATE: 2001-09-04
 NUMBER OF SEQ ID NOS: 31
 SEQ ID NO: 26
 LENGTH: 422
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-233-553-26

Query Match 60.3%; Score 35; DB 9; Length 422;
 Best Local Similarity 77.8%; Pred. No. 9.6e+02; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 2;

OY 1 LOEWLKKLK 9
 DB 73 LOEWLKKK 81

RESULT 32
 US-10-233-553-27
 Sequence 27, Application US/10233553
 Publication No. US20030125285A1

GENERAL INFORMATION:
 APPLICANT: NIPPON SHINYAKU CO., LTD.
 APPLICANT: HIRABAYASHI, Kazuko
 APPLICANT: YANO, Junichi
 TITLE OF INVENTION: Method of examining the efficacy of therapy with nucleic acid
 FILE REFERENCE: B-345
 CURRENT FILING DATE: 2002-12-19
 CURRENT APPLICATION NUMBER: JP 2001-267385
 PRIOR FILING DATE: 2001-09-04
 NUMBER OF SEQ ID NOS: 31
 SEQ ID NO: 27
 LENGTH: 422
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-233-553-27

US-10-233-553-27

Query Match 60.3%: Score 35; DB 9; Length 422;
Best Local Similarity 77.8%: Pred. No. 9.6e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKLK 9
|||||
DB 73 LOEWLKLK 81

RESULT 33

US-10-233-553-28
; Sequence 28, Application US/10233553
; Publication No. US20030125285A1
; GENERAL INFORMATION:
; APPLICANT: NIPPON SHINYAKU CO., LTD.
; APPLICANT: HIRABAYASHI, Kazuko
; TITLE OF INVENTION: Method of examining the efficacy of therapy with nucleic acid
; FILE REFERENCE: B-345
; CURRENT APPLICATION NUMBER: US/10/233,553
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: JP 2001-267385
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 31
; SEQ ID NO 28
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-233-553-28

Query Match 60.3%: Score 35; DB 9; Length 422;
Best Local Similarity 77.8%: Pred. No. 9.6e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKLK 9
|||||
DB 73 LOEWLKLK 81

RESULT 34

US-10-081-872-140
; Sequence 140, Application US/10081872
; Publication No. US20030125534A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Short, Jay M.
; APPLICANT: Machur, Eric J.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Krovacek, Janice S.
; APPLICANT: Slupsky, Margaret
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/081,872
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 140
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Environmental
US-10-081-872-140

Query Match 60.3%: Score 35; DB 9; Length 507;
Best Local Similarity 60.0%: Pred. No. 1.1e+03;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKLK 10
|||||
DB 265 LOEWLKLK 274

RESULT 35

US-10-211-962-85
; Sequence 85, Application US/10211962
; Publication No. US20030082640A1
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gottardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211,962
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US/09/562,737
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 85
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-211-962-85

Query Match 60.3%: Score 35; DB 9; Length 1024;
Best Local Similarity 63.6%: Pred. No. 2.2e+03;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKLK 11
|||||
DB 445 LOEWLKLK 455

RESULT 36

US-09-946-239-9
; Sequence 9, Application US/09946239
; Patent No. US20020044945A1
; GENERAL INFORMATION:
; APPLICANT: Bair, Margaret C.
; TITLE OF INVENTION: Polypeptide Sequences
; FILE REFERENCE: 18617, 0059
; CURRENT APPLICATION NUMBER: US/09/946,239
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 09/238,303, US 60/072,927
; PRIOR FILING DATE: 1999-01-28, 1998-01-29
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 9
; LENGTH: 1150
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: protein encoded by the pol gene of a recombinant viral
; OTHER INFORMATION: clone constructed from the genomic DNA of a Pallas's cat felin
; OTHER INFORMATION: Immunodeficiency virus
US-09-946-239-9

Query Match 60.3%: Score 35; DB 10; Length 1150;
Best Local Similarity 55.6%: Pred. No. 2.5e+03;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKLK 9
|||||
DB 1029 LKWKIKFR 1037

```
RESULT 37
US-10-223-070-21
; Sequence 21, Application US/10222070
; Publication No. US20030109045A1
; GENERAL INFORMATION:
; APPLICANT: NELSON, RICHARD S.
; APPLICANT: DING, XIN SHUN
; TITLE OF INVENTION: RNA SILENCING SUPPRESSION
; FILE REFERENCE: NBLF:0060S
; CURRENT APPLICATION NUMBER: US/10/223,070
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/313,185
; PRIOR FILING DATE: 2002-08-17
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 1718
; TYPE: PRT
; ORGANISM: Shallos virus X
US-10-223-070-21

Query Match      60.3%; Score 35; DB 9; Length 1718;
Best Local Similarity 55.6%; Pred. No. 3.7e+03;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      3  EWLKKLKKM 11
DB      1342 QWVKLEKV 1350

RESULT 38
US-09-736-968A-105
; Sequence 105, Application US/09736968A
; Patent No. US20020169283A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-7 Transmembrane Protein
; FILE REFERENCE: 020054-000611US
; CURRENT APPLICATION NUMBER: US/09/736,968A
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
```

```
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 105
; LENGTH: 2008
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human CLASP-4
US-09-736-968A-105

Query Match      60.3%; Score 35; DB 9; Length 2008;
Best Local Similarity 54.5%; Pred. No. 4.2e+03;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1  LQEWLKKLKKM 11
DB      195 MEWLITLTKI 205

RESULT 39
US-09-978-244A-29
; Sequence 29, Application US/0978244A
; Publication No. US2003010392A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter S.
; APPLICANT: Garman, Jonathan D.
; APPLICANT: Candia III, Albert F.
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP MEMBRANE PROTEINS
; FILE REFERENCE: 020554-000161US
; CURRENT APPLICATION NUMBER: US/09/978,244A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/310,028
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/737,246
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 09/736,969
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 09/736,960
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 09/736,968
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/240,545
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 2008
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: CLASP-4
US-09-978-244A-29

Query Match      60.3%; Score 35; DB 9; Length 2008;
Best Local Similarity 54.5%; Pred. No. 4.2e+03;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1  LQEWLKKLKKM 11
DB      195 MEWLITLTKI 205
```

```
RESULT 40
US-09-736-969A-2
; Sequence 2, Application US/09736969A
; Patent No. US20020068302A1
; GENERAL INFORMATION:
; APPLICANT: Garman, Jonathan David
; APPLICANT: Lu, Peter
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-4 Transmembrane Protein
; FILE REFERENCE: 020054-000411US
; CURRENT APPLICATION NUMBER: US/09/736, 969A
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2008
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: full length human CLASP-4
US-09-736-969A-2

Query Match      60.3%; Score 35; DB 10; Length 2008;
Best Local Similarity 54.5%; Pred. No. 4.2e+03;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 LQEWLKKLKKM 11
DB      195 MEWLITLTKI 205
```

```
RESULT 41
US-09-736-969A-91
; Sequence 91, Application US/09736969A
; Patent No. US20020068302A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-4 Transmembrane Protein
```

```
; FILE REFERENCE: 020054-000411US
; CURRENT APPLICATION NUMBER: US/09/736, 969A
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 91
; LENGTH: 2008
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human CLASP-4
US-09-736-969A-91

Query Match      60.3%; Score 35; DB 10; Length 2008;
Best Local Similarity 54.5%; Pred. No. 4.2e+03;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 LQEWLKKLKKM 11
DB      195 MEWLITLTKI 205
```

```
RESULT 42
US-09-736-960-88
; Sequence 88, Application US/09736960
; Patent No. US2002010267A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-5 Transmembrane Protein
; FILE REFERENCE: 020054-000511US
; CURRENT APPLICATION NUMBER: US/09/736, 960
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
```

```
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 88
; LENGTH: 2008
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human CLASP-4
; US-09-736-960-88
```

```
Query Match      60.3%; Score 35; DB 10; Length 2008;
Best Local Similarity 54.5%; Pred. No. 4.2e+03;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
Oy      1 LOEWLKLKLM 11
       :||| |||
Db      195 MEWLITLTKI 205
```

```
RESULT 43
US-09-978-244A-10
; Sequence 10, Application US/09978244A
; Publication No. US20030103992A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter S
; APPLICANT: Garman, Jonathan D.
; APPLICANT: Candia III, Albert F.
; TITLE OF INVENTION: CLASP MEMBRANE PROTEINS
; FILE REFERENCE: 020554-000161US
; CURRENT FILING DATE: US/09/978,244A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/310,028
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/737,246
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 09/736,969
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 09/736,960
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 09/736,968
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/240,545
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
```

```
; PRIOR FILING DATE: 2000-10-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 10
; LENGTH: 2073
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: CLASP-4
; US-09-978-244A-10
```

```
Query Match      60.3%; Score 35; DB 9; Length 2073;
Best Local Similarity 54.5%; Pred. No. 4.4e+03;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
Oy      1 LOEWLKLKLM 11
       :||| |||
Db      260 MEWLIMLTKI 270
```

```
RESULT 44
US-09-865-989-54
; Sequence 54, Application US/09865989
; Publication No. US2003000827A1
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Butner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/865,989
; FILING DATE: 25-MAY-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/465,719
; FILING DATE: 17-DEC-1999
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0006-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELETYPE: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. US2003000827A1e
; SEQUENCE DESCRIPTION: SEQ ID NO: 54:
; US-09-865-989-54

Query Match      58.6%; Score 34; DB 9; Length 22;
Best Local Similarity 60.0%; Pred. No. 80;
```

Matches 6: Conservative 2: Mismatches 2: Indels 0: Gaps 0:

Qy 1 LQEWLKKLKK 10
| | | | |
Db 10 LNEWLEALKQ 19

RESULT 45

US-10-099-574A-54
; Sequence 54, Application US/10099574A
; Publication No. US20030060604A1
; GENERAL INFORMATION:

APPLICANT: Dasseux, Jean-Louis
APPLICANT: Sekul, Renate
APPLICANT: Butner, Klaus
APPLICANT: Cornut, Isabelle
APPLICANT: Metz, Gunther
TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
NUMBER OF SEQUENCES: 254
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/099,574A
FILING DATE: 29-SEP-1997

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 009196-0005-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556

TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 54:

SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: NO. US20030060604A1e
US-10-099-574A-54

Query Match 58.6%; Score 34; DB 9; Length 22;
Best Local Similarity 60.0%; Pred. No. 80;
Matches 6: Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LQEWLKKLKK 10
| | | | |
Db 10 LNEWLEALKQ 19

RESULT 46

US-09-820-053A-9
; Sequence 9, Application US/09820053A
; Publication No. US20030083243A1
; GENERAL INFORMATION:

APPLICANT: Owen, Donald R.
TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES
FILE REFERENCE: HELX027

CURRENT APPLICATION NUMBER: US/09/820,053A
CURRENT FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 165
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9

LENGTH: 23
TYPE: PRT

ORGANISM: ARTIFICIAL SEQUENCE
FEATURE:
OTHER INFORMATION: SYNTHETIC SEQUENCE
NAME/KEY: MOD_RES
LOCATION: (23)

OTHER INFORMATION: AMIDATION
US-09-820-053A-9

Query Match 58.6%; Score 34; DB 9; Length 23;
Best Local Similarity 63.6%; Pred. No. 84;
Matches 7: Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LQEWLKKLKKM 11
| | | | |
Db 8 LKKLKKLKKL 18

RESULT 47

US-09-820-053A-161
; Sequence 161, Application US/09820053A
; Publication No. US20030083243A1
; GENERAL INFORMATION:

APPLICANT: Owen, Donald R.
TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES
FILE REFERENCE: HELX027
CURRENT APPLICATION NUMBER: US/09/820,053A
NUMBER OF SEQ ID NOS: 165

CURRENT FILING DATE: 2001-03-28
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 161
LENGTH: 23

TYPE: PRT
ORGANISM: ARTIFICIAL SEQUENCE
FEATURE:
OTHER INFORMATION: SYNTHETIC SEQUENCE

NAME/KEY: MOD_RES
LOCATION: (23)
OTHER INFORMATION: AMIDATION
US-09-820-053A-161

Query Match 58.6%; Score 34; DB 9; Length 23;
Best Local Similarity 63.6%; Pred. No. 84;
Matches 7: Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LQEWLKKLKKM 11
| | | | |
Db 8 LKKLKKLKKL 18

RESULT 48

US-10-109-171-9
; Sequence 9, Application US/10109171
; Publication No. US20030109452A1
; GENERAL INFORMATION:

APPLICANT: Owen, Donald R.
TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES AND METHODS FOR THEIR USE
FILE REFERENCE: HELX028
CURRENT APPLICATION NUMBER: US/10/109,171

CURRENT FILING DATE: 2002-03-28
NUMBER OF SEQ ID NOS: 165
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9

LENGTH: 23
TYPE: PRT
ORGANISM: ARTIFICIAL SEQUENCE
FEATURE:

OTHER INFORMATION: SYNTHETIC SEQUENCE
NAME/KEY: MOD_RES
LOCATION: (23)
OTHER INFORMATION: AMIDATION
US-10-109-171-9

Query Match 58.6%; Score 34; DB 9; Length 23;
Best Local Similarity 63.6%; Pred. No. 84;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LQEWLKKLKKM 11
|:|:|:|:|:|:
DB 8 LKKLKKLKKL 18

RESULT 49
US-10-109-171-161
Sequence 161, Application US/10109171
Publication No. US20030109452A1
GENERAL INFORMATION:
APPLICANT: Owen, Donald R.
TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES AND METHODS FOR THEIR USE
FILE REFERENCE: HELX028
CURRENT APPLICATION NUMBER: US/10/109.171
CURRENT FILING DATE: 2002-03-28
NUMBER OF SEQ ID NOS: 165
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 161
LENGTH: 23
TYPE: PPT
ORGANISM: ARTIFICIAL SEQUENCE
FEATURE:
OTHER INFORMATION: SYNTHETIC SEQUENCE
NAME/KEY: MOD_RES
LOCATION: (23)
OTHER INFORMATION: AMIDATION
US-10-109-171-161

Query Match 58.6%; Score 34; DB 9; Length 23;
Best Local Similarity 63.6%; Pred. No. 84;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LQEWLKKLKKM 11
|:|:|:|:|:|:
DB 8 LKKLKKLKKL 18

RESULT 50
US-10-097-079-70
Sequence 70, Application US/10097079
Patent No. US20020132973A1
GENERAL INFORMATION:
APPLICANT: Condon, Stephen M.
Morize, Isabelle
TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, Malistop 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/097.079
FILING DATE: 13-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/228,990
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/046,472
FILING DATE: 14-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Martin Esq., Michael B.
REGISTRATION NUMBER: 37,521
REFERENCE/DOCKET NUMBER: A2678B-WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-2793
TELEFAX: (610) 454-2798
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: NO. US20020132973A1 Relevant
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: Peptide
LOCATION: 8
OTHER INFORMATION: /product= "Nle"
FEATURE:
NAME/KEY: Peptide
LOCATION: 18
OTHER INFORMATION: /product= "Nle"
FEATURE:
NAME/KEY: Peptide
LOCATION: 25..29
OTHER INFORMATION: /product= "OTHER"
/note= "The side chains of Lys at position 25 and Asp at position 29 are linked by an amide bond."
FEATURE:
NAME/KEY: Peptide
LOCATION: 31
OTHER INFORMATION: /product= "OTHER"
/note= "This C-terminal amino acid is an amide, i.e., CONH2."
SEQUENCE DESCRIPTION: SEQ ID NO: 70:
US-10-097-079-70

Query Match 58.6%; Score 34; DB 12; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EWLKKL 8
|:|:|:|:|:
DB 22 EWLKKL 27

Search completed: July 11, 2003, 10:52:13
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 10:42:58 ; Search time 69 Seconds
(without alignments)
21.243 Million cell updates/sec

Title: US-09-828-592-10
Perfect score: 58
Sequence: 1 LQEWLKKLKKM 11

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 300 summaries

Database :

_A_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	100.0	11	23	AAU75185
2	41	70.7	11	23	AAU75186
3	41	70.7	212	22	ABG23248
4	40	69.0	11	23	AAU75183
5	40	69.0	108	22	ABG19285
6	40	69.0	139	21	AAG23368
7	40	69.0	175	21	AAG15566
8	40	69.0	175	21	AAG23367
9	40	69.0	177	21	AAG15565
10	40	69.0	177	21	AAG23366

11	40	69.0	209	21	AAG27746	Arabidopsis thalia
12	40	69.0	211	21	AAG27745	Arabidopsis thalia
13	40	69.0	303	21	AAG15564	Arabidopsis thalia
14	40	69.0	337	21	AAG27744	Arabidopsis thalia
15	40	69.0	432	19	AAW59840	Nature protein seq
16	40	69.0	464	4	AAW30445	Sequence of human
17	40	69.0	464	12	AAAR10381	Antithrombin III m
18	40	69.0	464	12	AAAR10383	Antithrombin III m
19	40	69.0	464	12	AAAR10384	Antithrombin III m
20	40	69.0	464	12	AAAR10390	Antithrombin III m
21	40	69.0	464	12	AAAR10380	Antithrombin III m
22	40	69.0	464	12	AAAR10382	Antithrombin III m
23	40	69.0	464	12	AAAR10385	Antithrombin III m
24	40	69.0	464	12	AAAR10386	Antithrombin III m
25	40	69.0	464	12	AAAR10387	Antithrombin III m
26	40	69.0	464	12	AAAR10388	Antithrombin III m
27	40	69.0	464	12	AAAR10389	Antithrombin III m
28	40	69.0	464	14	AAAR42895	Human antithrombin
29	40	69.0	464	14	AAAR42896	Human antithrombin
30	40	69.0	464	14	AAAR42897	Human antithrombin
31	40	69.0	464	14	AAAR42898	Human antithrombin
32	40	69.0	464	14	AAAR42899	Human antithrombin
33	40	69.0	464	14	AAAR42900	Human antithrombin
34	40	69.0	464	14	AAAR42901	Human antithrombin
35	40	69.0	464	14	AAAR42902	Human antithrombin
36	40	69.0	464	14	AAAR42903	Human antithrombin
37	40	69.0	464	14	AAAR42904	Human antithrombin
38	40	69.0	464	14	AAAR42905	Human antithrombin
39	40	69.0	464	14	AAAR42906	Human antithrombin
40	40	69.0	464	14	AAAR42907	Human antithrombin
41	40	69.0	464	14	AAAR42908	Human antithrombin
42	40	69.0	464	14	AAAR42909	Human antithrombin
43	40	69.0	464	14	AAAR42910	Human antithrombin
44	40	69.0	464	14	AAAR42911	Human antithrombin
45	40	69.0	464	14	AAAR42912	Human antithrombin
46	40	69.0	464	14	AAAR42913	Human antithrombin
47	40	69.0	464	14	AAAR42914	Human antithrombin
48	40	69.0	464	14	AAAR42915	Human antithrombin
49	40	69.0	464	14	AAAR42916	Human antithrombin
50	40	69.0	464	14	AAAR42917	Human antithrombin
51	40	69.0	464	14	AAAR42918	Human antithrombin
52	40	69.0	464	14	AAAR42919	Human antithrombin
53	40	69.0	464	14	AAAR42920	Human antithrombin
54	40	69.0	464	14	AAAR42921	Human antithrombin
55	40	69.0	464	14	AAAR42922	Human antithrombin
56	40	69.0	464	14	AAAR42923	Human antithrombin
57	40	69.0	464	14	AAAR42924	Human antithrombin
58	40	69.0	464	14	AAAR42925	Human antithrombin
59	40	69.0	464	14	AAAR42926	Human antithrombin
60	40	69.0	464	14	AAAR42927	Human antithrombin
61	40	69.0	464	14	AAAR42928	Human antithrombin
62	40	69.0	464	14	AAAR42929	Human antithrombin
63	40	69.0	464	14	AAAR42930	Human antithrombin
64	40	69.0	464	14	AAAR42931	Human antithrombin
65	40	69.0	464	14	AAAR42932	Human antithrombin
66	40	69.0	464	14	AAAR42933	Human antithrombin
67	40	69.0	464	14	AAAR42934	Human antithrombin
68	40	69.0	464	14	AAAR42935	Human antithrombin
69	40	69.0	464	14	AAAR42936	Human antithrombin
70	40	69.0	464	14	AAAR42937	Human antithrombin
71	40	69.0	464	14	AAAR42938	Human antithrombin
72	40	69.0	464	14	AAAR42939	Human antithrombin
73	40	69.0	464	14	AAAR42940	Human antithrombin
74	40	69.0	464	14	AAAR42941	Human antithrombin
75	40	69.0	464	21	AAI92224	Human antithrombin
76	40	69.0	464	22	AAAB37963	Human antithrombin
77	40	69.0	465	20	AAAY49551	Human antithrombin
78	40	69.0	707	22	ABG19287	Novel human diagno
79	40	69.0	934	23	AAE24153	Human 23479 protel
80	40	69.0	3353	23	AAU82706	Amino acid sequenc
81	39	67.2	10	23	ABU01660	158P107 related HL
82	39	67.2	85	21	AAAG20622	Arabidopsis thalia
83	39	67.2	131	21	AAAG35852	Arabidopsis thalia

84	39	67.2	155	21	AAG38604	Arabidopsis thalia	157	36	62.1	2441	19	AAW40058	Cellular transcrip
85	39	67.2	196	21	AAG35851	Arabidopsis thalia	158	36	62.1	2441	21	AAV94252	Mouse nuclear CREB
86	39	67.2	209	21	AAG35850	Arabidopsis thalia	159	36	62.1	2441	22	ABBA4555	Mouse wound healin
87	39	67.2	220	21	AAG38603	Arabidopsis thalia	160	36	62.1	2442	22	ABBA4556	Human wound healin
88	39	67.2	233	21	AAG38602	Arabidopsis thalia	161	36	62.1	3190	22	ABBA4634	Amino acid sequenc
89	39	67.2	336	21	AAV54098	Enzyme EPSM involv	162	36	62.1	3275	22	ABAB70437	Drosophila melanog
90	39	67.2	336	21	AAV43800	Amino acid sequenc	163	35	60.3	9	23	ABJ01520	158P1D7 related HL
91	39	67.2	546	23	AAW47876	ECM-related proteol	164	35	60.3	9	23	ABJ01601	158P1D7 related HL
92	39	67.2	798	23	ABJ01957	Human hypothetical	165	35	60.3	11	23	AAU75182	Amino acid sequenc
93	39	67.2	798	23	AAU79942	Human SIRT-like pr	166	35	60.3	11	23	AAU75187	Amino acid sequenc
94	39	67.2	823	19	AAW98507	H. pylori GHPD 122	167	35	60.3	16	14	AAK36611	Linker consensus s
95	39	67.2	841	23	ABJ01956	158P1D7 SSH protei	168	35	60.3	18	13	AAK20979	Sequence of amphip
96	39	67.2	841	23	AAE22222	Human toll like re	169	35	60.3	18	13	AAK22832	Amphiphilic peptid
97	39	67.2	841	23	ABBS3270	Human polypeptide	170	35	60.3	18	20	AAV19402	Lecithin:cholester
98	39	67.2	1290	17	AAK90583	Phospholipase C-ga	171	35	60.3	18	20	AAV19148	Lecithin:cholester
99	39	67.2	1640	23	ABBS4727	Lactococcus lactis	172	35	60.3	18	20	AAV18894	Lecithin:cholester
100	38	65.5	55	23	ABP26731	Streptococcus poly	173	35	60.3	18	20	AAV18631	Lecithin:cholester
101	38	65.5	120	22	AAK93639	Human protein sequ	174	35	60.3	20	17	AAW06885	Protein C inhibito
102	38	65.5	293	8	AAK70183	Sequence encoded b	175	35	60.3	20	17	AAK92270	Peptide hormone (M
103	38	65.5	409	16	AAK67652	Glucumylase. Sacc	176	35	60.3	32	6	AAK50470	Peptide hormone (M
104	38	65.5	551	7	AAK92442	Herbicidally activ	177	35	60.3	59	22	AAK70743	Human bone marrow
105	38	65.5	918	23	AAK60723	Sequence of amphi	178	35	60.3	59	22	AAK31051	Peptide #5068 enco
106	37	63.8	18	13	AAK21355	Amphiphilic peptid	179	35	60.3	87	22	AAO10639	Human polypeptide
107	37	63.8	18	13	AAK22839	Novel signal trans	180	35	60.3	89	22	ABBS3444	Peptide #950 enco
108	37	63.8	114	22	AAU17543	Listeria monocytog	181	35	60.3	205	22	ABBS69046	Drosophila melanog
109	37	63.8	117	22	ABBS50101	Arabidopsis thalia	182	35	60.3	206	22	ABBS2559	Escherichia coli p
110	37	63.8	144	21	AAK05467	Arabidopsis thalia	183	35	60.3	212	22	ABBS08265	Novel human diagno
111	37	63.8	144	21	AAK07169	Arabidopsis thalia	184	35	60.3	232	21	AAK45176	Gene 3 human secre
112	37	63.8	144	21	AAK48112	Arabidopsis thalia	185	35	60.3	263	22	ABG08262	Novel human diagno
113	37	63.8	144	21	AAK50109	Arabidopsis thalia	186	35	60.3	271	21	AAK53611	Arabidopsis thalia
114	37	63.8	186	22	AAK06076	Propionibacterium	187	35	60.3	273	21	AAK55808	Arabidopsis thalia
115	37	63.8	196	21	AAK05466	Arabidopsis thalia	188	35	60.3	273	21	AAK55808	Arabidopsis thalia
116	37	63.8	196	21	AAK07168	Arabidopsis thalia	189	35	60.3	290	21	AAK59078	Arabidopsis thalia
117	37	63.8	196	21	AAK48111	Arabidopsis thalia	190	35	60.3	290	21	AAK59297	Arabidopsis thalia
118	37	63.8	196	21	AAK50108	Arabidopsis thalia	191	35	60.3	291	21	AAK57409	Arabidopsis thalia
119	37	63.8	206	22	ABG13309	Novel human diagno	192	35	60.3	313	21	AAK53609	Arabidopsis thalia
120	37	63.8	207	22	ABG12035	Novel human diagno	193	35	60.3	313	21	AAK55807	Arabidopsis thalia
121	37	63.8	209	21	AAK05465	Arabidopsis thalia	194	35	60.3	326	21	AAK55806	Arabidopsis thalia
122	37	63.8	209	21	AAK07167	Arabidopsis thalia	195	35	60.3	327	23	ABP41158	Human ovarian anti
123	37	63.8	209	21	AAK48110	Arabidopsis thalia	196	35	60.3	334	21	AAK20597	Arabidopsis thalia
124	37	63.8	209	21	AAK50107	Arabidopsis thalia	197	35	60.3	334	21	AAK20597	Arabidopsis thalia
125	37	63.8	209	22	AAK030520	Novel human secret	198	35	60.3	355	21	AAK20596	Arabidopsis thalia
126	37	63.8	429	15	AAK91958	Peptide expressed	199	35	60.3	355	21	AAK50133	Arabidopsis thalia
127	37	63.8	429	15	AAK51887	DAPA synthase enco	200	35	60.3	357	20	AAV34114	Maize cyclopropane
128	37	63.8	429	22	AAK34505	E. coli cellulair p	201	35	60.3	381	21	AAK50132	Arabidopsis thalia
129	37	63.8	431	20	AAK73903	E. coli DAP aminot	202	35	60.3	382	21	AAK20595	Arabidopsis thalia
130	37	63.8	492	22	ABG08723	Novel human diagno	203	35	60.3	387	8	AAK70502	Human protein C in
131	37	63.8	546	22	AAK87604	Novel human diagno	204	35	60.3	387	23	AAU78358	Modified human pro
132	37	63.8	546	22	AAK23561	Novel human enzyme	205	35	60.3	387	23	AAU78359	Modified human pro
133	37	63.8	568	22	AAK07314	Novel central nerv	206	35	60.3	398	22	AAK96622	Putative P. abysssi
134	37	63.8	624	22	AAU22925	Novel human enzyme	207	35	60.3	406	23	AAU78357	Human protein C in
135	37	63.8	931	22	AAU51666	Propionibacterium	208	35	60.3	422	21	AAV94971	Human secreted pro
136	37	63.8	966	16	AAK76065	Yeast MSH2 protein	209	35	60.3	422	22	AAK39150	Human polypeptide
137	36	62.1	115	22	AAK000540	Human polypeptide	210	35	60.3	422	22	AAK60501	Human cell cycle a
138	36	62.1	146	23	ABBS5587	Lactococcus lactis	211	35	60.3	425	22	AAK29079	Novel human diagno
139	36	62.1	168	23	AAU17122	Novel signal trans	212	35	60.3	481	18	AAK21754	Novel 1-aminocycl
140	36	62.1	217	18	AAK40468	Bovine growth horm	213	35	60.3	491	23	ABBS2336	Herbicidally activ
141	36	62.1	217	18	AAK40468	Mouse neuronal reg	214	35	60.3	491	23	AAK40936	Human polypeptide
142	36	62.1	225	22	ABG10028	Novel human diagno	215	35	60.3	671	22	ABG29015	Novel human diagno
143	36	62.1	275	23	ABBA9384	Listeria monocytog	216	35	60.3	1024	23	ABBA4862	LDL receptor bindi
144	36	62.1	310	23	ABBS5124	Lactococcus lactis	217	35	60.3	1150	22	AAE11018	Recombinant viral
145	36	62.1	314	22	ABG25940	Novel human diagno	218	35	60.3	1150	23	ABAB08110	FIV-Omas pol prote
146	36	62.1	314	22	AAE06570	Schizandra chinens	219	35	60.3	2008	22	AAU04090	Human cadherin-lik
147	36	62.1	424	22	AAU34574	E. coli cellulair p	220	35	60.3	2008	23	ABG61703	Human cadherin-lik
148	36	62.1	428	22	AAU38258	S. salmonella typhi c	221	34	58.6	9	23	ABJ01423	158P1D7 related HL
149	36	62.1	551	22	AAK82287	S. epidermidis ope	222	34	58.6	9	23	ABJ01516	158P1D7 related HL
150	36	62.1	560	23	ABP38308	Staphylococcus epi	223	34	58.6	9	23	ABJ01987	158P1D7 related HL
151	36	62.1	575	23	ABG34135	Antibody productio	224	34	58.6	10	23	ABJ01733	158P1D7 related HL
152	36	62.1	2404	22	ABBS69209	Drosophila melanog	225	34	58.6	11	14	AAK45131	Amphiphilic peptid
153	36	62.1	2414	16	AAK48882	Transcription fact	226	34	58.6	11	14	AAK33972	Amphiphilic peptid
154	36	62.1	2414	19	AAK40057	Cellular transcrip	227	34	58.6	11	14	AAK31162	C-terminal substc.
155	36	62.1	2414	23	ABBS06340	Human p300 protein	228	34	58.6	11	14	AAK35380	Amphiphilic peptid
156	36	62.1	2441	16	AAK79054	CREB binding prote	229	34	58.6	11	15	AAK55986	Ion channel formin

230	34	58.6	11	15	AAR50562	Amphiphillic pepti
231	34	58.6	11	15	AAR50447	Amphiphillic peptid
232	34	58.6	11	15	AAR56947	Peptide which neut
233	34	58.6	11	15	AAR59064	Cancer treating, a
234	34	58.6	11	16	AAR90258	Ion-channel formin
235	34	58.6	11	17	AAR90906	Magainin-derived a
236	34	58.6	11	17	AAR99123	Amphiphillic peptid
237	34	58.6	11	19	AAM66297	Amphiphillic peptid
238	34	58.6	11	19	AAM66522	Peptide used to ma
239	34	58.6	11	20	AAV10749	Biologically activ
240	34	58.6	11	23	AAE22462	Ion-channel formin
241	34	58.6	12	16	AAR90272	Peptide used to ma
242	34	58.6	12	20	AAV10772	Biologically activ
243	34	58.6	12	23	AAE22485	Amphiphillic peptid
244	34	58.6	14	14	AAR45138	Amphiphillic peptid
245	34	58.6	14	14	AAR33974	C-terminal subst.
246	34	58.6	14	14	AAR31164	Amphiphillic peptid
247	34	58.6	14	14	AAR35382	Ion channel formin
248	34	58.6	14	15	AAR55988	Amphiphillic pepti
249	34	58.6	14	15	AAR50564	Amphiphillic peptid
250	34	58.6	14	15	AAR50449	Peptide which neut
251	34	58.6	14	15	AAR56949	Cancer treating, a
252	34	58.6	14	15	AAR59066	Ion-channel formin
253	34	58.6	14	16	AAR90265	Amphiphillic peptid
254	34	58.6	14	19	AAM66491	Peptide used to ma
255	34	58.6	14	20	AAV10756	Biologically activ
256	34	58.6	14	23	AAE22469	Sequence of amphip
257	34	58.6	15	13	AAR21403	Amphiphillic peptid
258	34	58.6	15	14	AAR45121	Amphiphillic peptid
259	34	58.6	15	14	AAR45122	Amphiphillic peptid
260	34	58.6	15	14	AAR45137	Amphiphillic peptid
261	34	58.6	15	14	AAR33962	Amphiphillic peptid
262	34	58.6	15	14	AAR33963	Amphiphillic peptid
263	34	58.6	15	14	AAR31152	C-terminal subst.
264	34	58.6	15	14	AAR31153	C-terminal subst.
265	34	58.6	15	14	AAR35370	Amphiphillic peptid
266	34	58.6	15	14	AAR35371	Amphiphillic peptid
267	34	58.6	15	15	AAR55976	Ion channel formin
268	34	58.6	15	15	AAR55977	Ion channel formin
269	34	58.6	15	15	AAR55991	Ion channel formin
270	34	58.6	15	15	AAR50437	Amphiphillic peptid
271	34	58.6	15	15	AAR50438	Amphiphillic peptid
272	34	58.6	15	15	AAR50552	Amphiphillic pepti
273	34	58.6	15	15	AAR50553	Amphiphillic pepti
274	34	58.6	15	15	AAR56937	Peptide which neut
275	34	58.6	15	15	AAR56938	Peptide which neut
276	34	58.6	15	15	AAR56952	Peptide which neut
277	34	58.6	15	15	AAR59069	Cancer treating, a
278	34	58.6	15	15	AAR59069	Cancer treating, a
279	34	58.6	15	15	AAR59054	Cancer treating, a
280	34	58.6	15	16	AAR90264	Ion-channel formin
281	34	58.6	15	16	AAR90143	Ion-channel formin
282	34	58.6	15	16	AAR90144	Ion-channel formin
283	34	58.6	15	17	AAW07257	Amphiphillic antim
284	34	58.6	15	19	AAM66486	Amphiphillic peptid
285	34	58.6	15	20	AAV10740	Peptide used to ma
286	34	58.6	15	20	AAV10755	Peptide used to ma
287	34	58.6	15	20	AAV10739	Peptide used to ma
288	34	58.6	15	23	AAE22452	Biologically activ
289	34	58.6	15	23	AAE22453	Biologically activ
290	34	58.6	15	23	AAE22468	Biologically activ
291	34	58.6	16	14	AAR45120	Amphiphillic peptid
292	34	58.6	16	14	AAR45120	Amphiphillic peptid
293	34	58.6	16	14	AAR33961	Amphiphillic ion ch
294	34	58.6	16	14	AAR33961	Amphiphillic peptid
295	34	58.6	16	14	AAR31151	C-terminal subst.
296	34	58.6	16	14	AAR35369	Amphiphillic peptid
297	34	58.6	16	14	AAR39079	Biologically activ
298	34	58.6	16	15	AAR55975	Ion channel formin
299	34	58.6	16	15	AAR50551	Amphiphillic pepti
300	34	58.6	16	15	AAR50436	Amphiphillic peptid
					AAR56936	Peptide which neut

ALIGNMENTS

```

RESULT 1
AAU75185
ID AAU75185 standard; peptide: 11 AA.
XX
AC AAU75185:
DT 08-MAY-2002 (first entry)
XX
DE Amino acid sequence of human antithrombin H-helix mutant AT-pos.
XX
KW Human; antithrombin: AT; H-helix; AT-pos; coagulation deficiency;
KW Humana; factor VIII; thrombin; T-TM; anticoagulant; hemostatic;
KW mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US2001055591-A1.
XX
PD 27-DEC-2001.
XX
PF 06-APR-2001; 2001US-0828592.
XX
PR 07-APR-2000; 2000US-195872P.
XX
PA (WALS/) WALSTON T.
PA (COOP/) COOPER S.
PA (REZA/) REZAI A.
XX
PI Walston T, Cooper S, Rezaie A;
XX
DR WPI: 2002-170988/22.
XX
PT Amino acid sequence useful for inhibiting thrombin activity comprises
PT antithrombin containing an H-helix having an amino acid modified to
PT have a more positive charge than an H-helix of non-modified
PT antithrombin
XX
PS Example 1: Page 5; 20pp: English.
XX
CC The present invention relates to mutant human antithrombin (AT) having
CC a modified H-helix to have a more positive charge (AT-pos) than a
CC H-helix of the wild type antithrombin. The amino acid sequences of
CC the modified AT H-helices are provided in the specification. The
CC mutant AT can be used for treating coagulation deficiency in a
CC patient, for treating haemophilia, and for extending the bioavailability
CC of factor VIII in a patient. The modified AT modulates the activity of
CC thrombin, and extends the length of time factor VIII is present in the
CC blood-stream of the patient and reduces the frequency and/or dosage of
CC factor VIII infusions needed by the patient. The mutant AT has enhanced
CC inhibitory activity against the bound form T-TM and thus blocks a
CC negative feedback loop, resulting in enhanced production of thrombin.
CC The present sequence represents the amino acid sequence of a human
CC AT H-helix mutant.
XX
SQ Sequence 11 AA:
XX
Query Match 100.0%; Score 58; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LQEWLKKLKKM 11
11
DB 1 LQEWLKKLKKM 11
11
RESULT 2
AAU75186
ID AAU75186 standard; peptide: 11 AA.
XX

```

AC AAU75186;
 XX 08-MAY-2002 (first entry)
 DE Amino acid sequence of human antithrombin H-helix mutant AT-neut.
 XX
 KW Human; antithrombin; AT; H-helix; AT-pos; coagulation deficiency;
 KW haemophilia; factor VIII; thrombin; T-TM; anticoagulant; haemostatic;
 KW mutant; mutlein.
 XX Homo sapiens.
 OS Synthetic.
 OS
 XX US2001055591-A1.
 PN
 XX 27-DEC-2001.
 PD
 XX 06-APR-2001; 2001US-0828592.
 PF
 XX 07-APR-2000; 2000US-195872P.
 PR
 XX (WALS/) WALSTON T.
 PA (COOP/) COOPER S.
 PA (REZA/) REZAIE A.
 XX
 PI Walston T, Cooper S, Rezaie A;
 XX
 DR WPI: 2002-170988/22.
 XX
 PT Amino acid sequence useful for inhibiting thrombin activity comprises
 PT antithrombin containing an H-helix having an amino acid modified to
 PT have a more positive charge than an H-helix of non-modified
 PT antithrombin -
 XX
 PS Example 1; Page 5; 20pp: English.
 XX
 CC The present invention relates to mutant human antithrombin (AT) having
 CC a modified H-helix to have a more positive charge (AT-pos) than a
 CC H-helix of the wild type antithrombin. The amino acid sequences of
 CC the modified AT H-helices are provided in the specification. The
 CC mutant AT can be used for treating coagulation deficiency in a
 CC patient, for treating haemophilia, and for extending the bioavailability
 CC of factor VIII in a patient. The modified AT modulates the activity of
 CC thrombin, and extends the length of time factor VIII is present in the
 CC blood-stream of the patient and reduces the frequency and/or dosage of
 CC factor VIII infusions needed by the patient. The mutant AT has enhanced
 CC inhibitory activity against the bound form T-TM and thus blocks a
 CC negative feedback loop, resulting in enhanced production of thrombin.
 CC The present sequence represents the amino acid sequence of a human
 CC AT H-helix mutant.
 CC
 XX
 SQ Sequence 11 AA;
 Query Match 70.7%; Score 41; DB 23; Length 11;
 Best Local Similarity 63.6%; Pred. No. 8.2;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LOEWLKKLKKM 11
 ||||| :||:|
 Db 1 LOEWLNQLQOM 11

KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US08631.
 PF
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PR
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 PI
 XX WPI: 2001-639362/73.
 DR N-PSDB; AAS87435.
 DR
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PT
 XX
 PS Claim 20; SEQ ID NO 53607; 103pp: English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABCG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 212 AA;
 Query Match 70.7%; Score 41; DB 22; Length 212;
 Best Local Similarity 54.5%; Pred. No. 1.4e+02;
 Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LOEWLKKLKKM 11
 |:||:|:||:|
 Db 111 LKEWVKRLKEL 121

RESULT 3
 ABG3248
 ID ABG3248 standard; Protein; 212 AA.
 AC
 XX ABG3248;
 XX
 DT 18-FEB-2002 (first entry)
 DT
 XX
 DE Novel human diagnostic protein #23239.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

RESULT 4
 AAU75183
 ID AAU75183 standard; peptide; 11 AA.
 AC
 XX AAU75183;
 XX
 DT 08-MAY-2002 (first entry)
 DT
 XX
 DE Amino acid sequence of human antithrombin H-helix (residues 304-314).
 XX
 KW Human; antithrombin; AT; H-helix; AT-pos; coagulation deficiency;
 KW haemophilia; factor VIII; thrombin; T-TM; anticoagulant; haemostatic.
 XX

OS Homo sapiens.
 XX
 PN US2001055591-A1.
 XX
 PD 27-DEC-2001.
 XX
 PF 06-APR-2001; 2001US-0828592.
 XX
 PR 07-APR-2000; 2000US-195872P.
 XX
 PA (WALS/) WALSTON T.
 PA (COOP/) COOPER S.
 PA (REZA/) REZAEI A.
 XX
 PI Walston T, Cooper S, Rezaei A;
 XX
 DR WPI: 2002-170988/22.
 XX
 PT Amino acid sequence useful for inhibiting thrombin activity comprises
 PT antithrombin containing an H-helix having an amino acid modified to
 PT have a more positive charge than an H-helix of non-modified
 PT antithrombin -
 XX
 PS Example 1; Page 5; 20pp; English.
 XX
 CC The present invention relates to mutant human antithrombin (AT) having
 CC a modified H-helix to have a more positive charge (AT-pos) than a
 CC H-helix of the wild type antithrombin. The amino acid sequences of
 CC the modified AT H-helices are provided in the specification. The
 CC mutant AT can be used for treating coagulation deficiency in a
 CC patient, for treating hemophilia, and for extending the bioavailability
 CC of factor VIII in a patient. The modified AT modulates the activity of
 CC thrombin, and extends the length of time factor VIII is present in the
 CC blood-stream of the patient and reduces the frequency and/or dosage of
 CC factor VIII infusions needed by the patient. The mutant AT has enhanced
 CC inhibitory activity against the bound form T-TM and thus blocks a
 CC negative feedback loop, resulting in enhanced production of thrombin.
 CC The present sequence represents the amino acid sequence of the wild
 CC type human AT H-helix (residues 304-314).
 CC
 SQ Sequence 11 AA:
 XX
 QY Query Match 69.0%; Score 40; DB 23; Length 11;
 XX Best Local Similarity 63.6%; Pred. No. 11;
 XX Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 DB 1 LOEWLKLKLM 11
 1 LOEWLDELLEM 11
 DB

PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 XX N-PSDB; AAS83472.
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID NO 49644; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 SQ Sequence 108 AA:
 XX
 QY Query Match 69.0%; Score 40; DB 22; Length 108;
 XX Best Local Similarity 63.6%; Pred. No. 1e+02;
 XX Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 DB 1 LOEWLKLKLM 11
 27 LOEWLDELLEM 37
 DB

RESULT 6
 AAG23368
 ID AAG23368 standard; Protein; 139 AA.
 XX
 AC AAG23368;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 26649.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000;
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 XX
 PR 05-MAR-1999; 99US-0123180.
 PR

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0138119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155486.
PR 23-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 69.0%; Score 40; DB 21; Length 139;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKLKK 10
:|||||
Db 1 MOEWLKKPK 10

RESULT 7
AAG15566

ID AAG15566 standard; Protein; 175 AA.

XX AAG15566;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 15868.

XX Protein identification: signal transduction pathway; metabolic pathway;
KW hydridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

OS EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 23-MAR-1999; 99US-0125748.

PR 25-MAR-1999; 99US-0126264.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145214.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.

PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 69.0%; Score 40; DB 21; Length 175;
Best Local Similarity 70.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LQEWLKLK 10
:||||| 1
Db 37 MQEWLKKYK 46

RESULT 8
AAG23367
ID AAG23367 standard; Protein; 175 AA.
XX

AC AAG23367;
XX

DT 17-OCT-2000 (first entry)
XX

DE Arabidopsis thaliana protein fragment SPQ ID NO: 26648.
XX

KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX

OS Arabidopsis thaliana.
XX

PN EP1033405-A2.
XX

PD 06-SEP-2000.
XX

PF 25-FEB-2000; 2000EP-0301439.
XX

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999;	99US-0132407.
PR 04-MAY-1999;	99US-0132484.
PR 05-MAY-1999;	99US-0132485.
PR 06-MAY-1999;	99US-0132486.
PR 07-MAY-1999;	99US-0132487.
PR 11-MAY-1999;	99US-0132863.
PR 14-MAY-1999;	99US-0134218.
PR 14-MAY-1999;	99US-0134219.
PR 14-MAY-1999;	99US-0134221.
PR 14-MAY-1999;	99US-0134370.
PR 18-MAY-1999;	99US-0134768.
PR 19-MAY-1999;	99US-0134941.
PR 20-MAY-1999;	99US-0135124.
PR 21-MAY-1999;	99US-0135353.
PR 24-MAY-1999;	99US-0135629.
PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 14-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139452.
PR 16-JUN-1999;	99US-0139453.
PR 17-JUN-1999;	99US-0139492.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0139459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 18-JUN-1999;	99US-0139763.
PR 21-JUN-1999;	99US-0139817.
PR 22-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140695.
PR 28-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140991.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141842.
PR 02-JUL-1999;	99US-0142154.
PR 02-JUL-1999;	99US-0142055.
PR 06-JUL-1999;	99US-0143390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0142977.
PR 13-JUL-1999;	99US-0143542.
PR 14-JUL-1999;	99US-0143624.
PR 15-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144085.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144332.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144352.
PR 20-JUL-1999;	99US-0144632.
PR 20-JUL-1999;	99US-0144884.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145085.
PR 22-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 27-JUL-1999;	99US-0145919.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 05-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 06-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149425.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 25-AUG-1999;	99US-0149930.
PR 26-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151067.
PR 30-AUG-1999;	99US-0151080.
PR 31-AUG-1999;	99US-0151303.
PR 01-SEP-1999;	99US-0151438.
PR 07-SEP-1999;	99US-0151930.
PR 10-SEP-1999;	99US-0153758.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155139.
PR 24-SEP-1999;	99US-0155486.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156458.
PR 04-OCT-1999;	99US-0156596.
PR 05-OCT-1999;	99US-0157117.
PR 06-OCT-1999;	99US-0157753.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0158293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 14-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match .69.0%; Score 40; DB 21; Length 175;
Best Local Similarity 70.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LOEWIKLK 10
DB 37 MOEWIKKYPK 46

RESULT 9
AAG15565
ID AAG15565 standard; Protein; 177 AA.

AC AAG15565;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 15867.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

OS EPI033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142800.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144088.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159328.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 69.0%; Score 40; DB 21; Length 177;
Best Local Similarity 70.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKLKK 10
:||||| |

DB 39 MOEWLKKYPK 48

RESULT 10
AAG23366

ID AAG23366 standard; Protein; 177 AA.

AC AAG23366;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 26647.

XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EP103405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131448.

PR 04-MAY-1999; 99US-0132408.

PR 05-MAY-1999; 99US-0132407.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999;	99US-0135353.
PR 24-MAY-1999;	99US-0135629.
PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 14-JUN-1999;	99US-0138847.
PR 16-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139452.
PR 17-JUN-1999;	99US-0139453.
PR 17-JUN-1999;	99US-0139492.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0139459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 18-JUN-1999;	99US-0139763.
PR 21-JUN-1999;	99US-0139817.
PR 22-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140695.
PR 28-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140991.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141842.
PR 01-JUL-1999;	99US-0142154.
PR 02-JUL-1999;	99US-0142055.
PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0142977.
PR 13-JUL-1999;	99US-0143542.
PR 14-JUL-1999;	99US-0143624.
PR 15-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144085.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144332.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144352.
PR 20-JUL-1999;	99US-0144632.
PR 20-JUL-1999;	99US-0144884.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145085.
PR 22-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 27-JUL-1999;	99US-0145919.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.

PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147483.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 23-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156556.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 18-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.
PR 21-OCT-1999;	99US-0160770.
PR 21-OCT-1999;	99US-0160814.
PR 22-OCT-1999;	99US-0160880.
PR 22-OCT-1999;	99US-0160980.
PR 22-OCT-1999;	99US-0160981.
PR 22-OCT-1999;	99US-0160989.
PR 25-OCT-1999;	99US-0161404.
PR 25-OCT-1999;	99US-0161405.
PR 25-OCT-1999;	99US-0161406.
PR 26-OCT-1999;	99US-0161359.
PR 26-OCT-1999;	99US-0161360.
PR 26-OCT-1999;	99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

Best Local Similarity 69.0%; Score 40; DB 21; Length 177;
Matches 7; Conservativity 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKKLKK 10
DB 39 MOEWLKKYK 48

RESULT 11

AAG27746
ID AAG27746 standard; Protein; 209 AA.

AC AAG27746;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 32704.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hydrolisation assay; genetic mapping; gene expression control; promoter;

XX Arabidopsis thaliana.

FN EPI033405-A2.

XX EPI033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 06-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134376.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 99US-0137528.

PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 02-JUL-1999; 99US-0141842.
PR 06-JUL-1999; 99US-0142055.
PR 09-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143277.
PR 14-JUL-1999; 99US-0143342.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147036.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 99US-0147192.
PR 99US-0147260.

PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149375.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

	Matches	7:	Conservative	1:	Mismatches	2:	Indels	0:	Gaps	0:
Oy		1	LOEWLKLKK	10.						
		:		1						
Db		37	MOEWLKKYK	46						
RESULT 12										
ID AAG27745										
AAG27745 standard; Protein: 211 AA.										
XX										
AC										
XX										
XX										
DT										
XX										
XX										
DE										
XX										
KW										
KM										
XX										
XX										
OS										
XX										
PN										
XX										
PD										
XX										
PF										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										

PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149422.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 69.0%; Score 40; DB 21; Length 211;
Best Local Similarity 70.0%; Pred. No. 2e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LOEWIKLKK 10
:|||||
Db 39 MOEWIKKPK 48

RESULT 13
AAG15564
ID AAG15564 standard; Protein: 303 AA.
XX AAG15564;
XX
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 15866.
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
KM hydrolisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144336.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147028.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156599.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158366.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161999.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162144.

Query Match 69.0%; Score 40; DB 21; Length 303;
Best Local Similarity 70.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKLKK 10
DB 165 MQEWLKKYK 174

RESULT 14
AAG27744
ID AAG27744 standard; Protein: 337 AA.

XX AC AAG27744;
XX AC AAG27744;
DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment spq ID NO: 32702.
DE Protein identification: signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS EPI033405-A2.
PN 06-SEP-2000.
PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
PF 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.

```
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140699.
PR 28-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141847.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143547.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144637.
PR 20-JUL-1999; 99US-0144884.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145291.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 69.0%; Score 40; DB 21; Length 337;
Best Local Similarity 70.0%; Pred. NO. 3.1e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LOEWLKKRK 10
Db 165 MQEWLKKRPK 174

RESULT 15
AAW59840
ID AAW59840 standard; Protein: 432 AA.
XX
AC AAW59840;
XX
DT 20-NOV-1998 (first entry)
XX
DE Mature protein sequence of antithrombin III (ATIII).
XX
KW Protein expression; monocytledon plant cell;
KW glycosylated alpha 1-antitrypsin; AAT; glycosylated antithrombin III;
KW ATIII; human serum albumin; HSA; subtilisin BPN'; treatment; emphysema;
KW antithrombotic; blood replacement.
```



```

XX OS Homo sapiens.
XX PN WO9836085-A1.
XX PD 20-AUG-1998.
XX PF 13-FEB-1998; 98WO-US03068.
XX PR 13-FEB-1997; 97US-0038170.
XX PR 13-FEB-1997; 97US-0037991.
XX PR 13-FEB-1997; 97US-0038168.
XX PR 13-FEB-1997; 97US-0038169.
XX PA (PHYT-) APPLIED PHYTOLOGICS INC.
XX PI Rodriguez RL, Suttiff TD;
XX DR WPI: 1998-467179/40.
XX N-PSDB: AAV41727.
XX PS Disclosure: Pages 29-30; 53pp; English.
XX CC The present sequence represents the mature protein of antithrombin III
XX CC (AIII). The protein is used to exemplify the invention. The
XX CC specification describes a method for producing mature heterologous
XX CC protein in monocotyledonous plant cells. The method comprises
XX CC transforming the cells with a chimeric gene comprising a monocotyledon
XX CC transcription regulator, inducible either during seed maturation or by
XX CC adding/removing a small molecule, DNA encoding the heterologous protein,
XX CC and DNA encoding a signal peptide, with the signal peptide causing
XX CC secretion of the protein from the cell. Proteins expressed in this
XX CC manner include mature glycosylated alpha 1-antitrypsin (AAT) with a
XX CC glycosylation pattern that significantly increases its serum half-life,
XX CC mature glycosylated antithrombin III (AIII), mature human serum albumin
XX CC (HSA) having the native folding pattern as shown by bilirubin-binding
XX CC characteristics, or mature active subtilisin BPN'. These proteins are
XX CC useful therapeutically (e.g. AAT for treating emphysema, AIII as
XX CC antithrombotic and HSA as blood replacement) or as industrial enzymes
XX CC (BPN' is used in detergents).
XX SQ Sequence 432 AA:

Query Match 69.0%; Score 40; DB 19; Length 432;
Best Local Similarity 63.6%; Pred. No. 3.9e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKKLKKM 11
DB 304 LOEWLDELDEM 314

RESULT 16
AAP30445
ID AAP30445 standard; Protein; 464 AA.
XX AC AAP30445;
XX DT 25-MAY-1992 (first entry)
XX DE Sequence of human antithrombin III (AIII) from cDNA clones pA62 and
XX DE pA68.
XX KW Thrombosis; therapy; cardiovascular disorders.
XX OS Homo sapiens.
XX FH Key
XX FT Peptide 1..32 Location/Qualifiers

```

```

FT XX /Label= signal
XX PN GB2116183-A.
XX PD 21-SEP-1983.
XX PF 02-MAR-1983; 83GB-0005786.
XX PR 30-JUL-1982; 82US-0403600.
XX PR 03-MAR-1982; 82GB-0006262.
XX PR 02-MAR-1983; 83GB-0005786.
XX PR 01-FEB-1985; 85US-0697178.
XX PA (GETH ) GENENTECH INC.
XX PI Bock SC, Lawn RM;
XX DR WPI: 1983-766797/38.
XX N-PSDB: AAN30203.
XX PS Disclosure: Fig 2; 14pp; English.
XX CC pA62 and pA68 are overlapping cDNA clones for AIII. pA62 is
XX CC incomplete at the 3' end and pA68 comprises 400 bp from the 3'
XX CC end, 84 bp 3' untranslated region and a poly(A) tail. Plasmids
XX CC pAIII-E7 and pAIII-J4 (claimed) contain a ligation product (pTA2)
XX CC of pA62 and pA68 (see AAN30203).
XX SQ Sequence 464 AA:

Query Match 69.0%; Score 40; DB 4; Length 464;
Best Local Similarity 63.6%; Pred. No. 4.2e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKKLKKM 11
DB 336 LOEWLDELDEM 346

RESULT 17
AAR10381
ID AAR10381 standard; Protein; 464 AA.
XX AC AAR10381;
XX DT 10-APR-1991 (first entry)
XX DE Antithrombin III mutant #2.
XX KW antithrombin 3; P-region variants; heparin-dependent; Factor Iia;
XX KW Factor Xa; melizothrombin; blood coagulation.
XX FH Key
XX FT Region 415..436
XX FT Location/Qualifiers
XX FT /Label= P-region
XX FT /note= "substitutions occur in this region which is
XX FT positions 383-404 in mature AIII"

WO9100291-A.
XX PD 10-JAN-1991.
XX PF 23-JUN-1990; 90WO-EP01026.
XX PR 26-JUN-1989; 89EP-0201675.
XX PA (ALKU ) AKZO NV.
XX PI Dijkema R, Visser A;

```

DR WPI; 1991-036710/05.
 XX Ser(394) to Ile and Asn(396) to Glu. (Numbers refer to mature
 CC Atrial). The substitutions result in a shift in the heparin-
 CC dependent inhibition profile of Atrial towards factors Ila and Xa.
 CC See also AAR10380-2 and AAR10384-R10390.
 CC XX
 PS Claim 2; Fig 2; 24pp: English.
 CC Mutations introduced into the Atrial reactive centre P-region are
 CC derived from the human prothrombin gene. Sequences known to be
 CC physiological substrates for Factor Xa were selected and used to
 CC replace the wild-type PvuII-StuI fragment of the original Atrial
 CC cDNA sequence (see AAQ10323). The substitutions are (given as
 CC positions in mature Atrial): Ala(384) to Glu,
 CC Ser(385) to Leu, Thr(386) to Leu, Ala(387) to Glu, Val(388) to Ser,
 CC Val(389) to Tyr and Ala(391) to Asp. The substitutions result in a
 CC shift in the heparin-dependent inhibition profile of Atrial towards
 CC factors Ila and Xa.
 CC See also AAR10380 and AAR10382-R10390.
 CC XX
 SQ Sequence 464 AA:
 Query Match 69.0%; Score 40; DB 12; Length 464;
 Best Local Similarity 63.6%; Pred. No. 4.2e+02;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 OY 1 LOEWLKLKLM 11
 ||||| :|::|
 Db 336 LOEWLDELEEM 346
 RESULT 18
 AAR10383
 ID AAR10383 standard; Protein; 464 AA.
 AC AAR10383;
 XX 10-APR-1991 (first entry)
 DT
 XX Antithrombin III mutant #4.
 DE
 XX Antithrombin III mutant #4.
 DE
 XX antithrombin 3; P-region variants; heparin-dependent; Factor Ila;
 KW Factor Xa; meizothrombin; blood coagulation.
 KM
 XX Key Location/Qualifiers
 FH 415..436
 FT /label= P-region
 FT /note= "substitutions occur in this region which is
 FT positions 383-404 in mature Atrial"
 FT
 XX WO9100291-A.
 PN 10-JAN-1991.
 PD
 XX 23-JUN-1990; 90MO-EP01026.
 PF
 XX 26-JUN-1989; 89EP-0201675.
 PR
 XX (ALKU) AKZO NV.
 PA
 PI DiJkema R, Visser A.
 XX WPI; 1991-036710/05.
 DR
 XX New modified antithrombin III variants - with altered
 PT heparin-dependent effect to factors Ila and Xa
 PT
 XX Claim 2; Fig 2; 24pp: English.
 PS
 CC Mutations introduced into the Atrial reactive centre P-region are
 CC derived from the human prothrombin gene. Sequences known to be
 CC physiological substrates for Factor Xa were selected and used to
 CC replace the wild-type PvuII-StuI fragment of the original Atrial
 CC cDNA sequence (see AAQ10323). The substitutions are (given as
 CC positions in mature Atrial): Ala(391) to Asp,
 CC and Ser(394) to Ile. The substitutions result in a shift in the
 CC heparin-dependent inhibition profile of Atrial towards factors Ila
 CC and Xa.
 CC See also AAR10380-3 and AAR10385-R10390.
 CC XX
 SQ Sequence 464 AA:
 Query Match 69.0%; Score 40; DB 12; Length 464;
 Best Local Similarity 63.6%; Pred. No. 4.2e+02;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 OY 1 LOEWLKLKLM 11
 ||||| :|::|

CC Ser(394) to Ile and Asn(396) to Glu. (Numbers refer to mature
 CC Atrial). The substitutions result in a shift in the heparin-
 CC dependent inhibition profile of Atrial towards factors Ila and Xa.
 CC See also AAR10380-2 and AAR10384-R10390.
 CC XX
 PS Claim 2; Fig 2; 24pp: English.
 CC Mutations introduced into the Atrial reactive centre P-region are
 CC derived from the human prothrombin gene. Sequences known to be
 CC physiological substrates for Factor Xa were selected and used to
 CC replace the wild-type PvuII-StuI fragment of the original Atrial
 CC cDNA sequence (see AAQ10323). The substitutions are (given as
 CC positions in mature Atrial): Ala(391) to Asp,
 CC and Ser(394) to Ile. The substitutions result in a shift in the
 CC heparin-dependent inhibition profile of Atrial towards factors Ila
 CC and Xa.
 CC See also AAR10380-3 and AAR10385-R10390.
 CC XX
 SQ Sequence 464 AA:
 Query Match 69.0%; Score 40; DB 12; Length 464;
 Best Local Similarity 63.6%; Pred. No. 4.2e+02;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 OY 1 LOEWLKLKLM 11
 ||||| :|::|

Db	336	LQEWLDELEEM	346
RESULT 20			
ID	AAR10390	standard; Protein; 464 AA.	
XX			
AC	AAR10390;		
XX			
DT	10-APR-1991	(first entry)	
XX			
DE	Antithrombin III P1' mutant.		
XX			
KW	antithrombin 3; P-region variants; heparin-dependent; Factor IIA;		
XX	Factor Xa; meizothrombin; blood coagulation.		
XX			
EH	Key	Location/Qualifiers	
FT	Region	415..436	
FT	/label=	P-region	
FT	/note=	"substitution occurs in this region, numbered 383-404 in mature ATIII, at the P1' site (=position 426)"	
FT	Misc-difference	426..426	
FT	/label=	Phe, Tyr, Leu, Val, Met, Thr, Gly, Ala	
FT	/note=	"P1' site"	
XX			
PN	W09100291-A.		
XX			
PD	10-JAN-1991.		
XX			
FE	23-JUN-1990;	90WO-EP01026.	
XX			
PR	26-JUN-1989;	89EP-0201675.	
XX			
PA	(ALKU) AKZO NV.		
XX			
PI	Dijkema R, Visser A;		
XX			
DR	WPI; 1991-036710/05.		
XX			
PT	New modified antithrombin III variants - with altered		
PT	heparin-dependent effect to factors IIA and Xa		
XX			
PS	Example; Fig 5; 24pp. English.		
XX			
CC	Ser at position 394 of mature ATIII (=426 in this sequence) is		
CC	substituted. The invention covers the substitution of Ile at this		
CC	position (= Mutant #10, see AAR10389). The nature of the substitution		
CC	was found to be crucial in conferring inhibitory profile to the		
CC	ATIII polypeptide. Substitutions (other than Ile) that represent a		
CC	hydrophilic character predict a specificity of heparin-dependent		
CC	inhibition towards Factor Xa.		
XX	See also AAO10323 and AAR10380-9.		
SO	Sequence	464 AA;	
Query Match		69.0%;	Score 40;
Best Local Similarity		63.6%;	DB 12;
Matches	7;	Conservative	3;
		Mismatches	1;
		Indels	0;
		Gaps	0;
OY	1	LOEWLKKLKM	11
		::::	
Db	336	LQEWLDELEEM	346
RESULT 21			
ID	AAR10380	standard; Protein; 464 AA.	
XX			
AC	AAR10380;		
XX			
DT	10-APR-1991	(first entry)	
XX			

```

DE Antithrombin III mutant #1.
XX
XX antithrombin 3; P-region variants; heparin-dependent; Factor IIA;
KW Factor Xa; meizothrombin; blood coagulation.
XX
FH Key location/Qualifiers
FT Region 415..436
FT /label= "P-region
FT /note- "Substitutions occur in this region which is
FT positions 383-404 of mature ATIII)"
XX
XX PN WO9100291-A.
XX PD 10-JAN-1991.
XX
XX PF 23-JUN-1990; 90WO-EPO1026.
XX PR 26-JUN-1989; 89EP-0201675.
XX (ALKU ) AKZO NV.
XX PA Dijkema R, Visser A;
XX PI WPI: 1991-036710/05.
XX DR
XX PT New modified antithrombin III variants - with altered
XX heparin-dependent effect to factors IIA and Xa
XX PS Claim 2; Fig 2; 24pp; English.
CC Mutations introduced into the ATIII reactive centre P-region are
CC derived from the human prothrombin gene. Sequences known to be
CC physiological substrates for Factor Xa were selected and used to
CC replace the wild-type PuuII-StuI fragment of the original ATIII
CC CDNA sequence (see AAQ10323). (Numbers are for mature ATIII).
CC The substitutions are Ala(384) to Glu, Ser(385) to Leu,
CC Thr(386) to Leu, Ala(387) to Glu, Val(388) to Ser, Val(389) to
CC Tyr, Ala(391) to Asp, Ser(394) to Ile, Leu(395) to Val, Asn(396) to
CC CC Gly, Pro(397) to Gly, Asn(398) to Ser, Arg(399) to Asp, Val(400) to
CC Ala, Thr(401) to Glu, Phe(402) to Ile and Lys(403) to Gly. The
CC substitutions result in a shift in the heparin-dependent inhibition
CC profile of ATIII towards factors IIA and Xa.
CC See also AAR10381-R10390.
XX
SQ Sequence 464 AA;
Query Match 69.0%; Score 40; DB 12; Length 464;
Best Local Similarity 63.6%; Pred. NO. 4.2e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
OY 1 LOEWLKLKLM 11
DB 336 LOEWMLDELLEM 346
RESULT 22
AARI0382
XX AARI0382 standard; Protein; 464 AA.
XX
XX AARI0382;
XX
XX 10-APR-1991 (first entry)
XX
XX DE Antithrombin III mutant #3.
XX
XX KW antithrombin 3; P-region variants; heparin-dependent; Factor IIA;
XX Factor Xa; meizothrombin; blood coagulation.
XX
XX FH Key Location/Qualifiers
XX FT Region 415..436
XX /label= "P-region
XX /note= "Substitutions occur in this region which is
XX positions 383-404 in mature ATIII"

```

```

XX  W09100291-A.
PN  10-JAN-1991.
XX  23-JUN-1990; 90WO-EP01026.
PD  26-JUN-1989; 89EP-0201675.
XX  (ALKU ) AKZO NV.
XX  Dijkema R, Visser A;
XX  WPI; 1991-036710/05.
XX  New modified antithrombin III variants - with altered
PT  heparin-dependent effect to factors IIA and Xa
XX  Claim 2; Fig 2; 24pp; English.
XX  Mutations introduced into the ATIII reactive centre P-region are
CC  derived from the human prothrombin gene. Sequences known to be
CC  physiological substrates for Factor Xa were selected and used to
CC  replace the wild-type PvuII-StuI fragment of the original ATIII
CC  cDNA sequence (see AAQ10323). The substitutions are (given as
CC  positions in mature ATIII): Ala(391) to Asp,
CC  Ser(394) to Ile, Leu(395) to Val and Asn(396) to Glu. The
CC  substitutions result in a shift in the heparin-dependent inhibition
CC  profile of ATIII towards factors IIA and Xa.
CC  See also AAR10380-1 and AAR10383-R10390.
CC  XX
SQ  Sequence 464 AA;

Query Match 69.0%; Score 40; DB 12; Length 464;
Best Local Similarity 63.6%; Pred. No. 4.2e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY  1 LQEWLKLKLM 11
    |||||:|:|
DB  336 LQEWLDELEEM 346

RESULT 23
AAR10385
ID  AAR10385 standard; Protein: 464 AA.
XX  AAR10385;
AC  10-APR-1991 (first entry)
XX  Antithrombin III mutant #6.
DE  Antithrombin III mutant #6.
XX  antithrombin 3; P-region variants; heparin-dependent; Factor IIA;
KW  Factor Xa; melzothrombin; blood coagulation.
XX  Key
FH  Location/Qualifiers
FT  415..436
FT  /label= P-region
FT  /note= "substitutions occur in this region which is
    numbered 383-404 in mature ATIII"
XX  W09100291-A.
PN  10-JAN-1991.
XX  23-JUN-1990; 90WO-EP01026.
XX  26-JUN-1989; 89EP-0201675.
XX  (ALKU ) AKZO NV.
XX  Dijkema R, Visser A;
XX

```

```

DR  WPI; 1991-036710/05.
XX  New modified antithrombin III variants - with altered
PT  heparin-dependent effect to factors IIA and Xa
XX  Claim 2; Fig 2; 24pp; English.
XX  Mutations introduced into the ATIII reactive centre P-region are
CC  derived from the human prothrombin gene. Sequences known to be
CC  physiological substrates for Factor Xa were selected and used to
CC  replace the wild-type PvuII-StuI fragment of the original ATIII
CC  cDNA sequence (see AAQ10323). The substitutions (given as positions
CC  in mature ATIII) are Ser(394) to Ile and Asn(396) to Glu. The
CC  substitutions result in a shift in the heparin-dependent
CC  inhibition profile of ATIII towards factors IIA and Xa.
CC  See also AAR10380-4 and AAR10386-R10390.
CC  XX
SQ  Sequence 464 AA;

Query Match 69.0%; Score 40; DB 12; Length 464;
Best Local Similarity 63.6%; Pred. No. 4.2e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY  1 LQEWLKLKLM 11
    |||||:|:|
DB  336 LQEWLDELEEM 346

RESULT 24
AAR10386
ID  AAR10386 standard; Protein: 464 AA.
XX  AAR10386;
AC  10-APR-1991 (first entry)
XX  Antithrombin III mutant #7.
DE  Antithrombin III mutant #7.
XX  antithrombin 3; P-region variants; heparin-dependent; Factor IIA;
KW  Factor Xa; melzothrombin; blood coagulation.
XX  Key
FH  Location/Qualifiers
FT  415..436
FT  /label= P-region
FT  /note= "substitutions occur in this region which is
    positions 383-404 in mature ATIII"
XX  W09100291-A.
PN  10-JAN-1991.
XX  23-JUN-1990; 90WO-EP01026.
XX  26-JUN-1989; 89EP-0201675.
XX  (ALKU ) AKZO NV.
XX  Dijkema R, Visser A;
XX  WPI; 1991-036710/05.
XX  New modified antithrombin III variants - with altered
PT  heparin-dependent effect to factors IIA and Xa
XX  Claim 2; Fig 2; 24pp; English.
XX  Mutations introduced into the ATIII reactive centre P-region are
CC  derived from the human prothrombin gene. Sequences known to be
CC  physiological substrates for Factor Xa were selected and used to
CC  replace the wild-type PvuII-StuI fragment of the original ATIII
CC  cDNA sequence (see AAQ10323). The substitutions are Ala(391) to Asp
CC  and Asn(396) to Glu. (Numbers refer to mature ATIII). The
CC  substitutions result in a shift in the heparin-dependent inhibition

```



```

FT Region 415..436
FT /label= P-region
FT /note= "substitutions occur in this region which is
FT numbered 383-404 in mature AtrialI"
XX WO9100291-A.
XX 10-JAN-1991.
XX
XX 23-JUN-1990; 90WO-EP01026.
XX
XX 26-JUN-1989; 89EP-0201675.
XX
XX (ALKU ) AKZO NV.
XX
XX Dijkema R, Visser A;
XX
XX WPI; 1991-036710/05.
XX
XX New modified antithrombin III variants - with altered
XX heparin-dependent effect to factors Iia and Xa
XX
XX Claim 2; Fig 2; 24pp; English.
XX
XX Mutations introduced into the AtrialI reactive centre P-region are
XX derived from the human prothrombin gene. Sequences known to be
XX physiological substrates for Factor Xa were selected and used to
XX replace the wild-type PvuII-StuI fragment of the original AtrialI
XX cDNA sequence (see AAQ10323). The only substitution (given as
XX position in mature AtrialI) is Ser(394) to Ile. The
XX substitution results in a shift in the heparin-dependent
XX inhibition profile of AtrialI towards factors Iia and Xa.
XX See also AAR10380-8 and AAR10390.
XX
XX Sequence 464 AA:
XX
XX Query Match 69.0%; Score 40; DB 12; Length 464;
XX Best Local Similarity 63.6%; Pred. No. 4.2e+02;
XX Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
OY 1 LOEWLKKLKKM 11
DB 336 LOEWLDELEEM 346

RESULT 28
AAR42895
ID AAR42895 standard; Protein: 464 AA.
XX
XX AAR42895;
XX
XX 13-MAY-1994 (first entry)
XX
XX Human antithrombin III (wild-type).
XX
XX AT III; mutant; serine protease inhibitor; serpin; thrombosis;
XX blood coagulation; anticoagulant; site-directed mutagenesis;
XX heparin binding site.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..32
XX Protein 33..464
XX /note= "mature wild-type AT III"
XX
XX Region 43..46
XX /note= "mutants include at least 1 mutation in
XX either this or one of the other indicated
XX regions"
XX
XX Region 73..79
XX /note= "mutants include at least 1 mutation in
XX either this or one of the other indicated

```

```

FT Region 157..165
FT /label= regions"
FT /note= "mutants include at least 1 mutation in
FT either this or one of the other indicated
FT regions"
XX Region 416..430
XX /note= "mutants include at least 1 mutation in
XX either this or one of the other indicated
XX regions"
XX
XX EP568833-A.
XX
XX 10-NOV-1993.
XX
XX 08-APR-1993; 93EP-0105829.
XX
XX 10-APR-1992; 92JP-0090488.
XX
XX 22-FEB-1993; 93JP-0031855.
XX
XX (EISA ) EISAI CO LTD.
XX
XX Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
XX Yoshitake S;
XX
XX WPI; 1993-352985/45.
XX
XX N-PSDB; AAQ50415.
XX
XX New human antithrombin III mutants of high antithrombin activity
XX in absence of heparin - useful as anticoagulant for treating
XX thrombotic disease
XX
XX Claim 1; Page 122-124; 137pp; English.
XX
XX The invention covers mutant versions of wild-type human AT III
XX (AAR42895) which have at least 1 amino acid mutation in the 11-14,
XX 41-47, 125-133 and/or 384-398 regions of the mature protein. The
XX mutants have high antithrombin activity, even in the absence of
XX heparin and are useful as anticoagulants. Mutation of the heparin
XX binding site prevents interaction of the mutant AT III with
XX vascular endothelial cells; the mutants have a long in vivo
XX half-life and are resistant to inactivation by neutrophil elastase.
XX
XX Sequence 464 AA:
XX
XX Query Match 69.0%; Score 40; DB 14; Length 464;
XX Best Local Similarity 63.6%; Pred. No. 4.2e+02;
XX Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
OY 1 LOEWLKKLKKM 11
DB 336 LOEWLDELEEM 346

RESULT 29
AAR42896
ID AAR42896 standard; Protein: 464 AA.
XX
XX AAR42896;
XX
XX 13-MAY-1994 (first entry)
XX
XX Human antithrombin III mutant 1R.
XX
XX AT III; mutant; serine protease inhibitor; serpin; thrombosis;
XX blood coagulation; anticoagulant; site-directed mutagenesis;
XX heparin binding site.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..32
XX Protein 33..464

```

```

FT      Region      /note- "mature wild-type AT III"
FT      416..430
FT      /note- "amino acids 384-398 of mature protein:
FT      wild-type Gly392 is substituted by Pro"
XX      EP568833-A.
XX      10-NOV-1993.
XX      08-APR-1993; 93EP-0105829.
XX      10-APR-1992; 92JP-0090488.
XX      22-FEB-1993; 93JP-0031855.
XX      (EISA ) EISAI CO LTD.
XX      Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
XX      Yoshitake S;
XX      WPI: 1993-352985/45.
XX      New human antithrombin III mutants of high antithrombin activity
XX      in absence of heparin - useful as anticoagulant for treating
XX      thrombotic disease
XX      Claim 20; Page 39-40; 137pp; English.
XX      The invention covers mutant versions of wild-type human AT III
XX      (AAR42895) which have at least 1 amino acid mutation in the 11-14,
XX      41-47, 125-133 and/or 384-398 regions of the mature protein. The
XX      mutants have high antithrombin activity, even in the absence of
XX      heparin and are useful as anticoagulants. Mutation of the heparin
XX      binding site prevents interaction of the mutant AT III with
XX      vascular endothelial cells; the mutants have a long in vivo
XX      half-life and are resistant to inactivation by neutrophil elastase.
XX      Sequence 464 AA;
XX      Query Match 69.0%; Score 40; DB 14; Length 464;
XX      Best Local Similarity 63.6%; Pred. No. 4.2e+02;
XX      Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY      1 LQEWLKKLKKM 11
DB      336 LQEWLDELEEM 346
RESULT 30
AAR42897
ID      AAR42897 standard; Protein: 464 AA.
XX      AAR42897;
XX      13-MAY-1994 (first entry)
XX      Human antithrombin III mutant 5R.
XX      AT III; mutant; serine protease inhibitor; serpin; thrombosis;
XX      blood coagulation; anticoagulant; site-directed mutagenesis;
XX      heparin binding site.
XX      Homo sapiens.
XX      Key Location/Qualifiers
XX      FT Peptide 1..32
XX      FT /label- signal_peptide 33..464
XX      FT Protein /note- "mature wild-type AT III"
XX      FT Region 416..430
XX      FT /note- "amino acids 384-398 of mature protein;
XX      FT wild-type Ala391-Gly392 are substituted
XX      FT by Phe-Pro"
XX

```

```

PN      EP568833-A.
XX      10-NOV-1993.
XX      08-APR-1993; 93EP-0105829.
XX      10-APR-1992; 92JP-0090488.
XX      22-FEB-1993; 93JP-0031855.
XX      (EISA ) EISAI CO LTD.
XX      Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
XX      Yoshitake S;
XX      WPI: 1993-352985/45.
XX      New human antithrombin III mutants of high antithrombin activity
XX      in absence of heparin - useful as anticoagulant for treating
XX      thrombotic disease
XX      Claim 21; Page 41-42; 137pp; English.
XX      The invention covers mutant versions of wild-type human AT III
XX      (AAR42895) which have at least 1 amino acid mutation in the 11-14,
XX      41-47, 125-133 and/or 384-398 regions of the mature protein. The
XX      mutants have high antithrombin activity, even in the absence of
XX      heparin and are useful as anticoagulants. Mutation of the heparin
XX      binding site prevents interaction of the mutant AT III with
XX      vascular endothelial cells; the mutants have a long in vivo
XX      half-life and are resistant to inactivation by neutrophil elastase.
XX      Sequence 464 AA;
XX      Query Match 69.0%; Score 40; DB 14; Length 464;
XX      Best Local Similarity 63.6%; Pred. No. 4.2e+02;
XX      Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY      1 LQEWLKKLKKM 11
DB      336 LQEWLDELEEM 346
RESULT 31
AAR42898
ID      AAR42898 standard; Protein: 464 AA.
XX      AAR42898;
XX      13-MAY-1994 (first entry)
XX      Human antithrombin III mutant 26R.
XX      AT III; mutant; serine protease inhibitor; serpin; thrombosis;
XX      blood coagulation; anticoagulant; site-directed mutagenesis;
XX      heparin binding site.
XX      Homo sapiens.
XX      Key Location/Qualifiers
XX      FT Peptide 1..32
XX      FT /label- signal_peptide 33..464
XX      FT Protein /note- "mature wild-type AT III"
XX      FT Region 416..430
XX      FT /note- "amino acids 384-398 of mature protein;
XX      FT wild-type Ala391-Gly392 are substituted
XX      FT by Ile-Pro"
XX      EP568833-A.
XX      10-NOV-1993.
XX      08-APR-1993; 93EP-0105829.
XX

```

```

XX 10-APR-1992; 92JP-0090488.
PR 22-FEB-1993; 93JP-0031855.
XX
PA (EISA ) EISAI CO LTD.
XX
PI Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
XX Yoshitake S;
XX WPI: 1993-352985/45.
XX
PS New human antithrombin III mutants of high antithrombin activity
PT in absence of heparin - useful as anticoagulant for treating
PR thrombotic disease
XX
PS Example 3; Page 42-44; 137pp; English.
XX
CC The invention covers mutant versions of wild-type human AT III
CC (AAR42895) which have at least 1 amino acid mutation in the 11-14,
CC 41-47, 125-133 and/or 384-398 regions of the mature protein. The
CC mutants have high antithrombin activity, even in the absence of
CC heparin and are useful as anticoagulants. Mutation of the heparin
CC binding site prevents interaction of the mutant AT III with
CC vascular endothelial cells; the mutants have a long in vivo
CC half-life and are resistant to inactivation by neutrophil elastase.
XX
SQ Sequence 464 AA:
OY Query Match 69.0%; Score 40; DB 14; Length 464;
Best Local Similarity 63.6%; Pred. No. 4.2e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
DB 336 LOEWLDELEEM 346
OY 1 LOEWLKKLKKM 11
|||||:|:|:|
DB 336 LOEWLDELEEM 346
RESULT 32
AAR42899
ID AAR42899 standard; Protein: 464 AA.
XX
AC AAR42899;
XX
DT 13-MAY-1994 (first entry)
XX
DE Human antithrombin III mutant 27R.
XX
KW AT III; mutant; serine protease inhibitor; serpin; thrombosis;
KW blood coagulation; anticoagulant; site-directed mutagenesis;
KW heparin binding site.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..32
FT /label= signal_peptide
FT Protein 33..464
FT /note= "mature wild-type AT III"
FT Region 416..430
FT /note= "amino acids 384-398 of mature protein;
FT wild-type Ala391-Gly392 are substituted
FT by Gly-Pro"
XX
PN EP568833-A.
XX
PD 10-NOV-1993.
XX
PE 08-APR-1993; 93EP-0105829.
XX
PR 10-APR-1992; 92JP-0090488.
PR 22-FEB-1993; 93JP-0031855.
XX
PA (EISA ) EISAI CO LTD.

```

```

XX Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
PI Yoshitake S;
XX WPI: 1993-352985/45.
XX
PS New human antithrombin III mutants of high antithrombin activity
PT in absence of heparin - useful as anticoagulant for treating
PR thrombotic disease
XX
PS Example 3; Page 44-45; 137pp; English.
XX
CC The invention covers mutant versions of wild-type human AT III
CC (AAR42895) which have at least 1 amino acid mutation in the 11-14,
CC 41-47, 125-133 and/or 384-398 regions of the mature protein. The
CC mutants have high antithrombin activity, even in the absence of
CC heparin and are useful as anticoagulants. Mutation of the heparin
CC binding site prevents interaction of the mutant AT III with
CC vascular endothelial cells; the mutants have a long in vivo
CC half-life and are resistant to inactivation by neutrophil elastase.
XX
SQ Sequence 464 AA:
OY Query Match 69.0%; Score 40; DB 14; Length 464;
Best Local Similarity 63.6%; Pred. No. 4.2e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
DB 336 LOEWLDELEEM 346
OY 1 LOEWLKKLKKM 11
|||||:|:|:|
DB 336 LOEWLDELEEM 346
RESULT 33
AAR42900
ID AAR42900 standard; Protein: 464 AA.
XX
AC AAR42900;
XX
DT 13-MAY-1994 (first entry)
XX
DE Human antithrombin III mutant 28R.
XX
KW AT III; mutant; serine protease inhibitor; serpin; thrombosis;
KW blood coagulation; anticoagulant; site-directed mutagenesis;
KW heparin binding site.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..32
FT /label= signal_peptide
FT Protein 33..464
FT /note= "mature wild-type AT III"
FT Region 416..430
FT /note= "amino acids 384-398 of mature protein;
FT wild-type Ala391-Gly392 are substituted
FT by Tyr-Pro"
XX
PN EP568833-A.
XX
PD 10-NOV-1993.
XX
PE 08-APR-1993; 93EP-0105829.
XX
PR 10-APR-1992; 92JP-0090488.
PR 22-FEB-1993; 93JP-0031855.
XX
PA (EISA ) EISAI CO LTD.
XX
PI Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
PI Yoshitake S;
XX WPI: 1993-352985/45.

```


XX New human antithrombin III mutants of high antithrombin activity
 PT in absence of heparin - useful as anticoagulant for treating
 PT thrombotic disease
 PS Example 3: Page 46-47; 137pp; English.
 CC The invention covers mutant versions of wild-type human AT III
 CC (AAR42895) which have at least 1 amino acid mutation in the 11-14,
 CC 41-47, 125-133 and/or 384-398 regions of the mature protein. The
 CC mutants have high antithrombin activity, even in the absence of
 CC heparin and are useful as anticoagulants. Mutation of the heparin
 CC binding site prevents interaction of the mutant AT III with
 CC vascular endothelial cells; the mutants have a long in vivo
 CC half-life and are resistant to inactivation by neutrophil elastase.
 CC
 XX Sequence 464 AA:
 SQ
 OY Query Match 69.0%; Score 40; DB 14; Length 464;
 Best Local Similarity 63.6%; Pred. No. 4.2e+02;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 DB 336 LOEWLDELEEM 346
 RESULT 34
 AAR42901
 ID AAR42901 standard; Protein: 464 AA.
 AC AAR42901;
 XX 13-MAY-1994 (first entry)
 DT Human antithrombin III mutant 29R.
 DE
 XX AT III; mutant; serine protease inhibitor; serpin; thrombosis;
 KW blood coagulation; anticoagulant; site-directed mutagenesis;
 KW heparin binding site.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..32
 FT /label= signal_peptide
 FT Protein 33..464
 FT /note= "mature wild-type AT III"
 FT Region 416..430
 FT /note= "amino acids 384-398 of mature protein;
 FT wild-type Ala391-Gly392 are substituted
 FT by Trp-Pro"
 FT
 FT
 FT
 XX EP568833-A.
 PN 10-NOV-1993.
 PD 10-NOV-1993.
 XX
 PD 08-APR-1993; 93EP-0105829.
 PE
 XX 10-APR-1992; 92JP-0090488.
 PR 22-FEB-1993; 93JP-0031855.
 PR
 XX (EISA) EISAI CO LTD.
 PA
 XX Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
 PI Yoshitake S;
 XX WPI: 1993-352985/45.
 DR
 XX New human antithrombin III mutants of high antithrombin activity
 PT in absence of heparin - useful as anticoagulant for treating
 PT thrombotic disease
 PT
 XX

PS Example 3: Page 47-48; 137pp; English.
 XX The invention covers mutant versions of wild-type human AT III
 CC (AAR42895) which have at least 1 amino acid mutation in the 11-14,
 CC 41-47, 125-133 and/or 384-398 regions of the mature protein. The
 CC mutants have high antithrombin activity, even in the absence of
 CC heparin and are useful as anticoagulants. Mutation of the heparin
 CC binding site prevents interaction of the mutant AT III with
 CC vascular endothelial cells; the mutants have a long in vivo
 CC half-life and are resistant to inactivation by neutrophil elastase.
 CC
 XX Sequence 464 AA:
 SQ
 OY Query Match 69.0%; Score 40; DB 14; Length 464;
 Best Local Similarity 63.6%; Pred. No. 4.2e+02;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 DB 336 LOEWLDELEEM 346
 RESULT 35
 AAR42902
 ID AAR42902 standard; Protein: 464 AA.
 AC AAR42902;
 XX 13-MAY-1994 (first entry)
 DT Human antithrombin III mutant 30R.
 DE
 XX AT III; mutant; serine protease inhibitor; serpin; thrombosis;
 KW blood coagulation; anticoagulant; site-directed mutagenesis;
 KW heparin binding site.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..32
 FT /label= signal_peptide
 FT Protein 33..464
 FT /note= "mature wild-type AT III"
 FT Region 416..430
 FT /note= "amino acids 384-398 of mature protein;
 FT wild-type Ala391-Gly392 are substituted
 FT by Val-Pro"
 FT
 FT
 FT
 XX EP568833-A.
 PN 10-NOV-1993.
 PD 10-NOV-1993.
 XX
 PD 08-APR-1993; 93EP-0105829.
 PE
 XX 10-APR-1992; 92JP-0090488.
 PR 22-FEB-1993; 93JP-0031855.
 PR
 XX (EISA) EISAI CO LTD.
 PA
 XX Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
 PI Yoshitake S;
 XX WPI: 1993-352985/45.
 DR
 XX New human antithrombin III mutants of high antithrombin activity
 PT in absence of heparin - useful as anticoagulant for treating
 PT thrombotic disease
 PT
 XX Example 3: Page 49-50; 137pp; English.
 CC The invention covers mutant versions of wild-type human AT III
 CC (AAR42895) which have at least 1 amino acid mutation in the 11-14,
 CC 41-47, 125-133 and/or 384-398 regions of the mature protein. The

CC mutants have high antithrombin activity, even in the absence of
CC heparin and are useful as anticoagulants. Mutation of the heparin
CC binding site prevents interaction of the mutant AT III with
CC vascular endothelial cells; the mutants have a long in vivo
CC half-life and are resistant to inactivation by neutrophil elastase.
XX
SQ Sequence 464 AA;

Query Match 69.0%; Score 40; DB 14; Length 464;
Best Local Similarity 63.6%; Pred. No. 4.2e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKLRLKRM 11
|||||:|:|
DB 336 LOEWLDELLEEM 346

RESULT 36
AAR42903
ID AAR42903 standard; Protein: 464 AA.

AC AAR42903;

DT 13-MAY-1994 (first entry)

DE Human antithrombin III mutant 46R.

KW AT III; mutant; serine protease inhibitor; serpin; thrombosis;
KW blood coagulation; anticoagulant; site-directed mutagenesis;
KW heparin binding site.

OS Homo sapiens.

XX Key Location/Qualifiers
FH Peptide 1..32
FT /label= signal_peptide
FT 33..464
FT Protein
FT /note= "mature wild-type AT III"
FT 416..430
FT Region
FT /note= "amino acids 384-398 of mature protein;
FT wild-type Ala391-Gly392 are substituted
FT by Leu-Pro"

XX EP568833-A.

XX 10-NOV-1993.

XX 08-APR-1993; 93EP-0105829.

XX 10-APR-1992; 92JP-0090488.

XX 22-FEB-1993; 93JP-0031855.

XX (EISA) EISAI CO LTD.

XX Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
PI Yoshitake S;
DR WPI: 1993-352985/45.

XX New human antithrombin III mutants of high antithrombin activity
PT in absence of heparin - useful as anticoagulant for treating
PT thrombotic disease
XX
XX Example 3; Page 50-52; 137pp; English.

CC The invention covers mutant versions of wild-type human AT III
CC (AAR42895) which have at least 1 amino acid mutation in the 11-14,
CC 41-47, 125-133 and/or 384-398 regions of the mature protein. The
CC mutants have high antithrombin activity, even in the absence of
CC heparin and are useful as anticoagulants. Mutation of the heparin
CC binding site prevents interaction of the mutant AT III with
CC vascular endothelial cells; the mutants have a long in vivo
CC half-life and are resistant to inactivation by neutrophil elastase.

XX
SQ Sequence 464 AA;

Query Match 69.0%; Score 40; DB 14; Length 464;
Best Local Similarity 63.6%; Pred. No. 4.2e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKLRLKRM 11
|||||:|:|
DB 336 LOEWLDELLEEM 346

RESULT 37
AAR42904
ID AAR42904 standard; Protein: 464 AA.

AC AAR42904;

DT 13-MAY-1994 (first entry)

DE Human antithrombin III mutant 39R.

XX AT III; mutant; serine protease inhibitor; serpin; thrombosis;
KW blood coagulation; anticoagulant; site-directed mutagenesis;
KW heparin binding site.

OS Homo sapiens.

XX Key Location/Qualifiers
FH Peptide 1..32
FT /label= signal_peptide
FT 33..464
FT Protein
FT /note= "mature wild-type AT III"
FT 416..430
FT Region
FT /note= "amino acids 384-398 of mature protein;
FT wild-type Ile390-Ala391-Gly392 are substid.
FT by Ala-Val-Pro"

XX EP568833-A.

XX 10-NOV-1993.

XX 08-APR-1993; 93EP-0105829.

XX 10-APR-1992; 92JP-0090488.

XX 22-FEB-1993; 93JP-0031855.

XX (EISA) EISAI CO LTD.

XX Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
PI Yoshitake S;
DR WPI: 1993-352985/45.

XX New human antithrombin III mutants of high antithrombin activity
PT in absence of heparin - useful as anticoagulant for treating
PT thrombotic disease
XX
XX Example 3; Page 52-53; 137pp; English.

CC The invention covers mutant versions of wild-type human AT III
CC (AAR42895) which have at least 1 amino acid mutation in the 11-14,
CC 41-47, 125-133 and/or 384-398 regions of the mature protein. The
CC mutants have high antithrombin activity, even in the absence of
CC heparin and are useful as anticoagulants. Mutation of the heparin
CC binding site prevents interaction of the mutant AT III with
CC vascular endothelial cells; the mutants have a long in vivo
CC half-life and are resistant to inactivation by neutrophil elastase.

SQ Sequence 464 AA;

Query Match 69.0%; Score 40; DB 14; Length 464;
Best Local Similarity 63.6%; Pred. No. 4.2e+02;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKKLKKM 11
 ||||| :|:|
 Db 336 LOEWLDELEEM 346

RESULT 38

AAR42905
 ID AAR42905 standard; Protein; 464 AA.

AC AAR42905;

DE 13-MAY-1994 (first entry)

DE Human antithrombin III mutant 40R.

KW AT III; mutant; serine protease inhibitor; serpin; thrombos-
 KW blood coagulation; anticoagulant; site-directed mutagenesis;
 KW heparin binding site.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Peptide 1..32 /label= signal_peptide

FT Protein 33..464 /note= "mature wild-type AT III"

FT Region 416..430 /note= "amino acids 384-398 of mature protein;
 FT wild-type Ile390-Ala391-Gly392 are substid.
 FT by Leu-Phe-Pro"

FT EP568833-A.

PD 10-NOV-1993.

PF 08-APR-1993; 93EP-0105829.

PR 10-APR-1992; 92JP-0090488.

PR 22-FEB-1993; 93JP-0031855.

PA (EISA) EISAI CO LTD.

PI Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
 PI Yoshitake S;

DR WPI; 1993-352985/45.

XX New human antithrombin III mutants of high antithrombin activity

PT in absence of heparin - useful as anticoagulant for treating

PT thrombotic disease

PS Example 3; Page 52-53; 137pp; English.

XX The invention covers mutant versions of wild-type human AT III

CC (AAR42895) which have at least 1 amino acid mutation in the 11-14,

CC 41-47, 125-133 and/or 384-398 regions of the mature protein. The

CC mutants have high antithrombin activity, even in the absence of

CC heparin and are useful as anticoagulants. Mutation of the heparin

CC binding site prevents interaction of the mutant AT III with

CC vascular endothelial cells; the mutants have a long in vivo

CC half-life and are resistant to inactivation by neutrophil elastase.

XX Sequence 464 AA;

Query Match 69.0%; Score 40; DB 14; Length 464;

Best Local Similarity 63.6%; Pred. No. 4.2e+02;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKKLKKM 11
 ||||| :|:|
 Db 336 LOEWLDELEEM 346

RESULT 39

AAR42906
 ID AAR42906 standard; Protein; 464 AA.

AC AAR42906;

DE 13-MAY-1994 (first entry)

DE Human antithrombin III mutant 48R.

KW AT III; mutant; serine protease inhibitor; serpin; thrombos-
 KW blood coagulation; anticoagulant; site-directed mutagenesis;
 KW heparin binding site.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Peptide 1..32 /label= signal_peptide

FT Protein 33..464 /note= "mature wild-type AT III"

FT Region 416..430 /note= "amino acids 384-398 of mature protein;
 FT wild-type Ile390-Ala391-Gly392 are substid.
 FT by Ala-Tyr-Pro"

FT EP568833-A.

PD 10-NOV-1993.

PF 08-APR-1993; 93EP-0105829.

PR 10-APR-1992; 92JP-0090488.

PR 22-FEB-1993; 93JP-0031855.

PA (EISA) EISAI CO LTD.

PI Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
 PI Yoshitake S;

DR WPI; 1993-352985/45.

XX New human antithrombin III mutants of high antithrombin activity

PT in absence of heparin - useful as anticoagulant for treating

PT thrombotic disease

PS Example 3; Page 55-56; 137pp; English.

XX The invention covers mutant versions of wild-type human AT III

CC (AAR42895) which have at least 1 amino acid mutation in the 11-14,

CC 41-47, 125-133 and/or 384-398 regions of the mature protein. The

CC mutants have high antithrombin activity, even in the absence of

CC heparin and are useful as anticoagulants. Mutation of the heparin

CC binding site prevents interaction of the mutant AT III with

CC vascular endothelial cells; the mutants have a long in vivo

CC half-life and are resistant to inactivation by neutrophil elastase.

XX Sequence 464 AA;

Query Match 69.0%; Score 40; DB 14; Length 464;

Best Local Similarity 63.6%; Pred. No. 4.2e+02;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKKLKKM 11
 ||||| :|:|
 Db 336 LOEWLDELEEM 346

RESULT 40

AAR42907
 ID AAR42907 standard; Protein; 464 AA.

xx	Homo sapiens.	
os		
xx		
fh	Key	Location/Qualifiers
ft	Peptide	1..32
ft	Protein	/label= signal_peptide
ft	Region	33..464
ft		/note= "mature wild-type AT III"
ft		416..430
ft		/note= "amino acids 384-398 of mature protein;
ft		wild-type Ile390-Ala391 are substituted
ft		by Ala-Val"
pn	EP568833-A.	
xx		
pd	10-NOV-1993.	
xx		
pe	08-APR-1993;	93EP-0105829.
xx		
pr	10-APR-1992;	92JP-0090488.
pr	22-FEB-1993;	93JP-0031855.
xx		
pa	(EISA) EISAI CO LTD.	
xx		
pi	Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;	
pl	Yoshitake S;	
xx		
dr	WPI; 1993-352985/45.	
xx		
pt	New human antithrombin III mutants of high antithrombin activity	
pt	in absence of heparin - useful as anticoagulant for treating	
pt	thrombotic disease	
xx		
ps	Example 3: Page 60-61; 137pp; English.	
xx		
cc	The invention covers mutant versions of wild-type human AT III	
cc	(AAR42895) which have at least 1 amino acid mutation in the 11-14,	
cc	41-47, 125-133 and/or 384-398 regions of the mature protein. The	
cc	mutants have high antithrombin activity, even in the absence of	
cc	heparin and are useful as anticoagulants. Mutation of the heparin	
cc	binding site prevents interaction of the mutant AT III with	
cc	vascular endothelial cells; the mutants have a long in vivo	
cc	half-life and are resistant to inactivation by neutrophil elastase.	
xx		
so	Sequence 464 AA:	
	Query Match 69.0%; Score 40; DB 14; Length 464;	
	Best Local Similarity 63.6%; Pred. NO. 4.2e+02;	
	Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0.	
oy	1 LQEWLKKLKKM 11	
	:::	
db	336 LQEWLDELEEM 346	
	RESULT 43	
	AAR42910	
	ID AAR42910 standard; Protein; 464 AA.	
xx		
ac	AAR42910;	
xx		
dt	13-MAY-1994 (first entry)	
xx		
de	Human antithrombin III mutant 34R.	
xx		
kw	AT III; mutant; serine protease inhibitor; serpin; thrombosis;	
kw	blood coagulation; anticoagulant; site-directed mutagenesis;	
xx	heparin binding site.	
os		
os	Homo sapiens.	
xx		
fh	Key	Location/Qualifiers
ft	Peptide	1..32

```

FT      Protein          /label= signal_peptide  

FT      Region          /note= "mature wild-type AT IIR"  

FT      /note= "amino acids 384-398 of mature protein;  

FT      wild-type Ile390-Ala391 are substituted  

FT      by Ala-Ile"  

XX      EP568833-A.  

XX      PD              10-NOV-1993.  

XX      PE              08-APR-1993;   93EP-0105829.  

XX      PR              10-APR-1992;   92JP-0090488.  

XX      PR              22-FEB-1993;   93JP-0031855.  

PA      (EISA ) EISAI CO LTD.  

XX      PI              Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;  

PI              Yoshitake S;  

DR      WPI; 1993-352985/45.  

XX      PT              New human antithrombin III mutants of high antithrombin activity  

PT              in absence of heparin - useful as anticoagulant for treating  

PT              thrombotic disease  

XX      PS              Example 3; Page 61-63; 137pp; English.  

XX      CC              The invention covers mutant versions of wild-type human AT III  

CC              (AAR42895) which have at least 1 amino acid mutation in the 11-14,  

CC              41-47, 125-133 and/or 384-398 regions of the mature protein. The  

CC              mutants have high antithrombin activity, even in the absence of  

CC              heparin and are useful as anticoagulants. Mutation of the heparin  

CC              binding site prevents interaction of the mutant AT III with  

CC              vascular endothelial cells; the mutants have a long in vivo  

CC              half-life and are resistant to inactivation by neutrophil elastase.  

SQ      Sequence       464 AA;  

QY      Query Match    69.0%; Score 40; DB 14; Length 464;  

        Best Local Similarity 63.6%; Pred No. 4.2e+02;  

        Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  

        QY      1 LQEWLKKLKRM 11  

        |||||::|:  

DB      336 LQEWMLDELEEM 346  

RESULT 44  

AAID AAR42911 standard; Protein: 464 AA.  

AC AAR42911;  

XX DT 13-MAY-1994 (first entry)  

XX DE Human antithrombin III mutant 35R.  

KW AT IIR; mutant; serine protease inhibitor; serpin; thrombosus;  

KW blood coagulation; anticoagulant; site-directed mutagenesis;  

KW heparin binding site.  

OS Homo sapiens.  

XX FH Key  

FH FT Peptide  

FT /label= signal_peptide  

FT /note= "mature wild-type AT IIR"  

FT /note= "amino acids 384-398 of mature protein;  

FT wild-type Ile390-Ala391 are substituted  

FT by Ala-Ile"  

FT Region  

FT /note= "amino acids 384-398 of mature protein;"
```

FT FT wild-type Ile390-Ala391 are substituted
 by Ala-Leu"
 XX XX
 PD EP568833-A.
 10-NOV-1993.
 XX XX
 PF 08-APR-1993; 93EP-0105829.
 XX XX
 PR 10-APR-1992; 92JP-0090488.
 PR 22-FEB-1993; 93JP-0031855.
 XX XX
 PA (EISA) EISAI CO LTD.
 XX XX
 PI Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
 PI Yoshitake S;
 DR WPI; 1993-352985/45.
 XX XX
 PT New human antithrombin III mutants of high antithrombin activity
 PT in absence of heparin - useful as anticoagulant for treating
 PT thrombotic disease
 PS Claim 22; Page 63-64; 137pp; English.
 XX XX
 CC The invention covers mutant versions of wild-type human AT III
 CC (AAR42895) which have at least 1 amino acid mutation in the 11-14,
 CC 41-47, 125-133 and/or 384-398 regions of the mature protein. The
 CC mutants have high antithrombin activity, even in the absence of
 CC heparin and are useful as anticoagulants. Mutation of the heparin
 CC binding site prevents interaction of the mutant AT III with
 CC vascular endothelial cells; the mutants have a long in vivo
 CC half-life and are resistant to inactivation by neutrophil elastase.
 CC XX
 SQ Sequence 464 AA;
 Query Match 69.0%; Score 40; DB 14; Length 464;
 Best Local Similarity 63.6%; Pred. No. 4.2e+02;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LOEWLKKLKKM 11
 ||||| : : : :
 DB 336 LOEWLDELEEM 346
 RESULT 45
 AAR42912
 ID AAR42912 standard; Protein: 464 AA.
 AC AAR42912;
 XX XX
 DT 13-MAY-1994 (first entry)
 XX XX
 DE Human antithrombin III mutant 38R.
 XX XX
 KW AT III; mutant; serine protease inhibitor; serpin; thrombosis;
 KW blood coagulation; anticoagulant; site-directed mutagenesis;
 KW heparin binding site.
 XX XX
 OS Homo sapiens.
 XX XX
 FH Key
 FT Peptide
 FT /label= signal_peptide
 FT Protein
 FT /note= "mature wild-type AT III"
 FT Region
 FT 416..430
 FT /note= "amino acids 384-398 of mature protein;
 FT wild-type Ile390-Ala391 are substituted
 FT by Gly-Ile"
 XX XX
 EP568833-A.

PD 10-NOV-1993.
 XX XX
 PF 08-APR-1993; 93EP-0105829.
 XX XX
 PR 10-APR-1992; 92JP-0090488.
 PR 22-FEB-1993; 93JP-0031855.
 XX XX
 PA (EISA) EISAI CO LTD.
 XX XX
 PI Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
 PI Yoshitake S;
 DR WPI; 1993-352985/45.
 XX XX
 PT New human antithrombin III mutants of high antithrombin activity
 PT in absence of heparin - useful as anticoagulant for treating
 PT thrombotic disease
 PS Example 3; Page 65-66; 137pp; English.
 XX XX
 CC The invention covers mutant versions of wild-type human AT III
 CC (AAR42895) which have at least 1 amino acid mutation in the 11-14,
 CC 41-47, 125-133 and/or 384-398 regions of the mature protein. The
 CC mutants have high antithrombin activity, even in the absence of
 CC heparin and are useful as anticoagulants. Mutation of the heparin
 CC binding site prevents interaction of the mutant AT III with
 CC vascular endothelial cells; the mutants have a long in vivo
 CC half-life and are resistant to inactivation by neutrophil elastase.
 CC XX
 SQ Sequence 464 AA;
 Query Match 69.0%; Score 40; DB 14; Length 464;
 Best Local Similarity 63.6%; Pred. No. 4.2e+02;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LOEWLKKLKKM 11
 ||||| : : : :
 DB 336 LOEWLDELEEM 346
 RESULT 46
 AAR42913
 ID AAR42913 standard; Protein: 464 AA.
 AC AAR42913;
 XX XX
 DT 13-MAY-1994 (first entry)
 XX XX
 DE Human antithrombin III mutant 9R.
 XX XX
 KW AT III; mutant; serine protease inhibitor; serpin; thrombosis;
 KW blood coagulation; anticoagulant; site-directed mutagenesis;
 KW heparin binding site.
 XX XX
 OS Homo sapiens.
 XX XX
 FH Key
 FT Peptide
 FT /label= signal_peptide
 FT Protein
 FT /note= "mature wild-type AT III"
 FT Region
 FT 416..430
 FT /note= "amino acids 384-398 of mature protein;
 FT wild-type Ala384 is substituted by Gly"
 XX XX
 PD EP568833-A.
 10-NOV-1993.
 XX XX
 PF 08-APR-1993; 93EP-0105829.
 XX XX
 PR 10-APR-1992; 92JP-0090488.
 PR 22-FEB-1993; 93JP-0031855.

```

XX PA (EISA ) EISAI CO LTD.
XX PI Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
XX PI Yoshitake S;
XX DR WPI: 1993-352985/45.
XX
PT New human antithrombin III mutants of high antithrombin activity
PT in absence of heparin - useful as anticoagulant for treating
PT thrombotic disease
XX
PS Example 3: Page 66-68; 137pp; English.
XX
CC The invention covers mutant versions of wild-type human AT III
CC (AAR42895) which have at least 1 amino acid mutation in the 11-14,
CC 41-47, 125-133 and/or 384-398 regions of the mature protein. The
CC mutants have high antithrombin activity, even in the absence of
CC heparin and are useful as anticoagulants. Mutation of the heparin
CC binding site prevents interaction of the mutant AT III with
CC vascular endothelial cells; the mutants have a long in vivo
CC half-life and are resistant to inactivation by neutrophil elastase.
XX
SQ Sequence 464 AA:
Query Match 69.0%; Score 40; DB 14; Length 464;
Best Local Similarity 63.6%; Pred. No. 4.2e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKLKLM 11
DB 336 LOEWLDELLEM 346

RESULT 47
AAR42914
ID AAR42914 standard; Protein: 464 AA.
XX
AC AAR42914;
XX
DT 13-MAY-1994 (first entry)
XX
DE Human antithrombin III mutant 19R.
XX
KW AT III; mutant; serine protease inhibitor; serpin; thrombosis;
KW blood coagulation; anticoagulant; site-directed mutagenesis;
KW heparin binding site.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..32
FT /label= signal_peptide
FT Protein 33..464
FT /note= "mature wild-type AT III"
FT Region 416..430
FT /note= "amino acids 384-398 of mature protein;
FT wild-type Val389 is substituted by Pro"
XX
FN EP568833-A.
PD 10-NOV-1993.
XX
PF 08-APR-1993; 93EP-0105829.
XX
PR 10-APR-1992; 92JP-0090488.
PR 22-FEB-1993; 93JP-0031855.
XX
PA (EISA ) EISAI CO LTD.
XX
PI Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
PI Yoshitake S;
XX

```

```

DR WPI: 1993-352985/45.
XX
PT New human antithrombin III mutants of high antithrombin activity
PT in absence of heparin - useful as anticoagulant for treating
PT thrombotic disease
XX
PS Example 3: Page 68-69; 137pp; English.
XX
CC The invention covers mutant versions of wild-type human AT III
CC (AAR42895) which have at least 1 amino acid mutation in the 11-14,
CC 41-47, 125-133 and/or 384-398 regions of the mature protein. The
CC mutants have high antithrombin activity, even in the absence of
CC heparin and are useful as anticoagulants. Mutation of the heparin
CC binding site prevents interaction of the mutant AT III with
CC vascular endothelial cells; the mutants have a long in vivo
CC half-life and are resistant to inactivation by neutrophil elastase.
XX
SQ Sequence 464 AA:
Query Match 69.0%; Score 40; DB 14; Length 464;
Best Local Similarity 63.6%; Pred. No. 4.2e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKLKLM 11
DB 336 LOEWLDELLEM 346

RESULT 48
AAR42915
ID AAR42915 standard; Protein: 464 AA.
XX
AC AAR42915;
XX
DT 13-MAY-1994 (first entry)
XX
DE Human antithrombin III mutant 24R.
XX
KW AT III; mutant; serine protease inhibitor; serpin; thrombosis;
KW blood coagulation; anticoagulant; site-directed mutagenesis;
KW heparin binding site.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..32
FT /label= signal_peptide
FT Protein 33..464
FT /note= "mature wild-type AT III"
FT Region 416..430
FT /note= "amino acids 384-398 of mature protein;
FT wild-type Ala387 is substituted by Phe"
XX
FN EP568833-A.
PD 10-NOV-1993.
XX
PF 08-APR-1993; 93EP-0105829.
XX
PR 10-APR-1992; 92JP-0090488.
PR 22-FEB-1993; 93JP-0031855.
XX
PA (EISA ) EISAI CO LTD.
XX
PI Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;
PI Yoshitake S;
XX
DR WPI: 1993-352985/45.
XX
PT New human antithrombin III mutants of high antithrombin activity
PT in absence of heparin - useful as anticoagulant for treating
PT thrombotic disease
XX

```

PS Example 3: Page 69-71; 137pp; English.

XX The invention covers mutant versions of wild-type human AT III

CC (AAR42895) which have at least 1 amino acid mutation in the 11-14,

CC 41-47, 125-133 and/or 384-398 regions of the mature protein. The

CC mutants have high antithrombin activity, even in the absence of

CC heparin and are useful as anticoagulants. Mutation of the heparin

CC binding site prevents interaction of the mutant AT III with

CC vascular endothelial cells; the mutants have a long in vivo

CC half-life and are resistant to inactivation by neutrophil elastase.

XX

SO Sequence 464 AA:

OY Query Match 69.0%; Score 40; DB 14; Length 464;

DB Best Local Similarity 63.6%; Pred. No. 4.2e+02;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LQEWIKRLKKM 11

DB 336 LQEWLDELEEM 346

RESULT 49

AAR42916

ID AAR42916 standard; Protein; 464 AA.

XX

AC AAR42916;

XX

DT 13-MAY-1994 (first entry)

XX

DE Human antithrombin III mutant 2R'.

XX

KW AT III; mutant; serine protease inhibitor; serpin; thrombosis;

KW blood coagulation; anticoagulant; site-directed mutagenesis;

KW heparin binding site.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FT Peptide 1..32

FT Protein /label= signal_peptide

FT 33..464

FT /note= "mature wild-type AT III"

FT 416..430

FT Region /note= "amino acids 384-398 of mature protein;

FT wild-type Asn398 is substituted by Glu"

XX

PN EP568833-A.

XX

PD 10-NOV-1993.

XX

PF 08-APR-1993; 93EP-0105829.

XX

PR 10-APR-1992; 92JP-0090488.

PR 22-FEB-1993; 93JP-0031855.

XX

PA (EISA) EISAI CO LTD.

XX

PI Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;

PI Yoshitake S;

XX

DR WPI; 1993-352985/45.

XX

PT New human antithrombin III mutants of high antithrombin activity.

PT in absence of heparin - useful as anticoagulant for treating

PT thrombotic disease

XX

PS Example 3: Page 71-72; 137pp; English.

XX

CC The invention covers mutant versions of wild-type human AT III

CC (AAR42895) which have at least 1 amino acid mutation in the 11-14,

CC 41-47, 125-133 and/or 384-398 regions of the mature protein. The

CC mutants have high antithrombin activity, even in the absence of

CC heparin and are useful as anticoagulants. Mutation of the heparin

CC binding site prevents interaction of the mutant AT III with

CC vascular endothelial cells; the mutants have a long in vivo

CC half-life and are resistant to inactivation by neutrophil elastase.

CC

SO Sequence 464 AA;

CC heparin and are useful as anticoagulants. Mutation of the heparin

CC binding site prevents interaction of the mutant AT III with

CC vascular endothelial cells; the mutants have a long in vivo

CC half-life and are resistant to inactivation by neutrophil elastase.

XX

SO Sequence 464 AA:

OY Query Match 69.0%; Score 40; DB 14; Length 464;

DB Best Local Similarity 63.6%; Pred. No. 4.2e+02;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LQEWIKRLKKM 11

DB 336 LQEWLDELEEM 346

RESULT 50

AAR42917

ID AAR42917 standard; Protein; 464 AA.

XX

AC AAR42917;

XX

DT 13-MAY-1994 (first entry)

XX

DE Human antithrombin III mutant 5R'.

XX

KW AT III; mutant; serine protease inhibitor; serpin; thrombosis;

KW blood coagulation; anticoagulant; site-directed mutagenesis;

KW heparin binding site.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FT Peptide 1..32

FT Protein /label= signal_peptide

FT 33..464

FT /note= "mature wild-type AT III"

FT 416..430

FT Region /note= "amino acids 384-398 of mature protein;

FT wild-type Asn398 is substituted by Arg"

XX

PN EP568833-A.

XX

PD 10-NOV-1993.

XX

PF 08-APR-1993; 93EP-0105829.

XX

PR 10-APR-1992; 92JP-0090488.

PR 22-FEB-1993; 93JP-0031855.

XX

PA (EISA) EISAI CO LTD.

XX

PI Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki N, Suzuki S;

PI Yoshitake S;

XX

DR WPI; 1993-352985/45.

XX

PT New human antithrombin III mutants of high antithrombin activity

PT in absence of heparin - useful as anticoagulant for treating

PT thrombotic disease

XX

PS Example 3: Page 73-74; 137pp; English.

XX

CC The invention covers mutant versions of wild-type human AT III

CC (AAR42895) which have at least 1 amino acid mutation in the 11-14,

CC 41-47, 125-133 and/or 384-398 regions of the mature protein. The

CC mutants have high antithrombin activity, even in the absence of

CC heparin and are useful as anticoagulants. Mutation of the heparin

CC binding site prevents interaction of the mutant AT III with

CC vascular endothelial cells; the mutants have a long in vivo

CC half-life and are resistant to inactivation by neutrophil elastase.

CC

SO Sequence 464 AA;

Query Match 69.0%; Score 40; DB 14; Length 464;
 Best Local Similarity 63.6%; Pred. No. 4.2e+02;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWIKKIKKM 11
 |||||:|:|
 Db 336 LOEWLDELEEM 346

Search completed: July 11, 2003, 10:48:29
 Job time : 75 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 10:45:48 ; Search time 40 Seconds
(without alignments)
26.437 Million cell updates/sec

Title: US-09-828-592-10
Perfect score: 58
Sequence: 1 LQEMLKLRKM 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 300 summaries

Database : PIR_73:.*
1: p1r1:.*
2: p1r2:.*
3: p1r3:.*
4: p1r4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	72.4	118	2 G70478	hypothetical prote
2	41	70.7	429	2 E85585	7.8-diaminopelargo
3	41	70.7	429	2 D90735	7.8-diaminopelargo
4	40	69.0	315	2 AF3270	Integrase (Importe
5	40	69.0	457	2 F71172	hypothetical prote
6	40	69.0	464	1 KHHU3	antithrombin III p
7	40	69.0	524	2 S77851	probable aspartate
8	40	69.0	745	2 H85048	probable transposo
9	40	69.0	1120	2 T01863	hypothetical prote
10	40	69.0	1138	2 A82939	membrane nuclease
11	40	69.0	1323	1 RRMGSM	RNA-directed RNA p
12	39	67.2	209	2 H85062	hypothetical prote
13	39	67.2	314	2 T46279	hypothetical prote
14	39	67.2	454	2 H83977	hypothetical prote
15	39	67.2	480	2 D75053	aminotransferase B
16	39	67.2	629	2 T14776	hypothetical prote
17	39	67.2	823	2 F64526	adenine/cytosine D
18	39	67.2	928	2 AH0950	DNA polymerase I f
19	39	67.2	955	2 E84022	hypothetical prote
20	39	67.2	1114	2 B86423	hypothetical prote
21	39	67.2	1290	2 A31317	1-phosphatidylinos
22	39	67.2	1290	2 A36456	1-phosphatidylinos
23	39	67.2	1291	2 S00666	1-phosphatidylinos
24	39	67.2	1640	2 D86798	prophage p13 prote
25	38	65.5	98	1 OPCM2	acylphosphatase (E
26	38	65.5	131	2 G69442	transcription regu
27	38	65.5	143	2 C75173	hypothetical prote
28	38	65.5	252	2 A64492	toxin sensitivity
29	38	65.5	313	2 S61990	hypothetical prote

30	38	65.5	334	2 F86462	ATG1-like protein,
31	38	65.5	402	2 T12745	hypothetical prote
32	38	65.5	583	2 F88642	protein C54B4.3 (1
33	38	65.5	767	1 JU0474	glucan 1,4-alpha-9
34	38	65.5	778	1 ALBYG	glucan 1,4-alpha-9
35	38	65.5	786	2 H69980	single-strand DNA-
36	38	65.5	4131	2 T21085	hypothetical prote
37	37	63.8	1117	2 AG1740	B. subtilis genera
38	37	63.8	117	2 A11370	B. subtilis genera
39	37	63.8	165	2 B87702	ribosomal protein
40	37	63.8	170	2 C64366	hypothetical prote
41	37	63.8	209	2 T05164	hypothetical prote
42	37	63.8	218	2 F72272	conserved hypothet
43	37	63.8	252	2 C70380	hypothetical prote
44	37	63.8	415	2 AE2237	hypothetical prote
45	37	63.8	429	1 XNECDP	adenosylmethionine
46	37	63.8	445	2 T42366	L-aminoadipate-sem
47	37	63.8	550	2 A12249	exopolyposphatase
48	37	63.8	578	2 D82910	aspartyl-tRNA synt
49	37	63.8	751	2 T24458	hypothetical prote
50	37	63.8	852	2 AD3232	confugal transfer
51	37	63.8	852	2 F44020	hypothetical prote
52	37	63.8	852	2 T08502	ltdE protein - Ent
53	37	63.8	964	2 S57379	MSH2 protein - yea
54	37	63.8	981	2 T41029	probable aminoadip
55	37	63.8	1070	2 G84982	exodeoxyribonucia
56	37	63.8	1643	1 RRMGNV	genome polyprotein
57	36	62.1	135	2 F85291	Yle homolog, A. tu
58	36	62.1	142	2 G95329	probable dehydroge
59	36	62.1	146	2 H86903	hypothetical prote
60	36	62.1	161	2 A69732	PBSX prophage ORF
61	36	62.1	169	2 A83077	probable signal pe
62	36	62.1	198	2 G42148	GMP-binding protei
63	36	62.1	200	2 S73781	hypothetical prote
64	36	62.1	201	2 E71085	hypothetical prote
65	36	62.1	275	2 B90540	conserved hypothet
66	36	62.1	287	2 B90540	conserved hypothet
67	36	62.1	310	2 G86846	ABC transporter AT
68	36	62.1	374	2 S50891	2-nitropropane dio
69	36	62.1	414	2 S18962	FBF15 protein - St
70	36	62.1	424	1 STECYT	tyrosine-tRNA liga
71	36	62.1	424	2 B90922	tyrosine-tRNA synt
72	36	62.1	424	2 AE0693	tyrosyl-tRNA synt
73	36	62.1	428	2 F85770	tyrosine-tRNA synt
74	36	62.1	436	2 G87756	protein T01A.3 (1
75	36	62.1	458	2 A30775	dihydrotolipamide a
76	36	62.1	479	2 E97337	aspartyl/asparagin
77	36	62.1	492	2 A28616	M5 protein precurs
78	36	62.1	498	2 G90067	malate quinone oxi
79	36	62.1	611	2 T41563	hypothetical prote
80	36	62.1	632	2 T16128	hypothetical prote
81	36	62.1	691	2 S48390	hypothetical prote
82	36	62.1	726	2 A10147	hypothetical prote
83	36	62.1	805	2 E71719	probable ATP-depen
84	36	62.1	805	2 E97717	virB4 protein prec
85	36	62.1	1056	2 T28636	zinc-metalloprotein
86	36	62.1	1085	2 F66712	hypothetical prote
87	36	62.1	1170	2 AC3595	transcription repa
88	36	62.1	1245	2 S51255	transcription membra
89	36	62.1	2414	2 A54277	transcription adap
90	36	62.1	2440	2 S39162	CREB-binding prote
91	36	62.1	2441	2 S39161	CREB-binding prote
92	36	62.1	3190	2 T13828	CREB-binding prote
93	35	60.3	63	2 A72467	probable ribosomal
94	35	60.3	109	2 A98001	phosphoserine phos
95	35	60.3	115	2 F96923	probable transcrip
96	35	60.3	134	2 A81062	conserved hypothet
97	35	60.3	134	2 G81807	hypothetical prote
98	35	60.3	140	2 T13855	centromere protein
99	35	60.3	159	2 T16702	hypothetical prote
100	35	60.3	175	2 B97314	probable transcrip
101	35	60.3	201	2 D75101	hypothetical prote
102	35	60.3	231	2 S50460	hypothetical prote

ATG1-like protein,
hypothetical prote
protein C54B4.3 (1
glucan 1,4-alpha-9
glucan 1,4-alpha-9
single-strand DNA-
hypothetical prote
B. subtilis genera
B. subtilis genera
ribosomal protein
hypothetical prote
hypothetical prote
conserved hypothet
hypothetical prote
hypothetical prote
adenosylmethionine
L-aminoadipate-sem
exopolyposphatase
aspartyl-tRNA synt
hypothetical prote
confugal transfer
hypothetical prote
ltdE protein - Ent
MSH2 protein - yea
probable aminoadip
exodeoxyribonucia
genome polyprotein
Yle homolog, A. tu
probable dehydroge
hypothetical prote
PBSX prophage ORF
probable signal pe
GMP-binding protei
hypothetical prote
hypothetical prote
conserved hypothet
ABC transporter AT
2-nitropropane dio
FBF15 protein - St
tyrosine-tRNA liga
tyrosine-tRNA synt
tyrosyl-tRNA synt
tyrosine-tRNA synt
protein T01A.3 (1
dihydrotolipamide a
aspartyl/asparagin
M5 protein precurs
malate quinone oxi
hypothetical prote
hypothetical prote
hypothetical prote
probable ATP-depen
virB4 protein prec
zinc-metalloprotein
hypothetical prote
transcription repa
transcription membra
transcription adap
CREB-binding prote
CREB-binding prote
A72467
probable ribosomal
phosphoserine phos
probable transcrip
conserved hypothet
hypothetical prote
centromere protein
hypothetical prote
probable transcrip
hypothetical prote
hypothetical prote

103	35	60.3	259	2	H64211	guanylate kinase (
104	35	60.3	264	2	S52102	thioether S-methyl
105	35	60.3	292	2	B90558	hypothetical prote
106	35	60.3	299	2	G82221	transcription regu
107	35	60.3	313	2	F86295	T24D18.18 protein
108	35	60.3	332	2	A70228	exrd protein homol
109	35	60.3	336	2	E72389	hypothetical prote
110	35	60.3	344	2	B90176	conserved hypotet
111	35	60.3	398	2	A75128	probable transamin
112	35	60.3	406	2	A39339	protein C inhibitor
113	35	60.3	411	2	F90215	hypothetical prote
114	35	60.3	497	2	S78286	conserved hypotet
115	35	60.3	545	2	T43513	Hsk1 protein kinas
116	35	60.3	556	2	D70940	probable PPE prote
117	35	60.3	564	2	F90965	hypothetical prote
118	35	60.3	564	2	F85813	hypothetical prote
119	35	60.3	569	2	H64959	probable membrane
120	35	60.3	590	2	S77267	secreted protein M
121	35	60.3	605	2	A96660	protein F2K1.20 f
122	35	60.3	721	2	T46037	hypothetical prote
123	35	60.3	793	2	A82126	hypothetical prote
124	35	60.3	796	2	D97065	transketolase [imp
125	35	60.3	1036	2	T38734	hypothetical prote
126	35	60.3	1121	2	G82120	transcription regu
127	35	60.3	1574	1	A38454	myosin MYO2 - yeas
128	35	60.3	1718	1	JQ1734	genome polyprotein
129	35	60.3	2514	2	T37320	ataxia telangiecta
130	35	60.3	2619	2	T24588	hypothetical prote
131	34	58.6	76	2	T17872	hypothetical prote
132	34	58.6	84	2	JE0379	NADH2 dehydrogenas
133	34	58.6	113	2	E72104	conserved hypotet
134	34	58.6	113	2	F86518	hypothetical prote
135	34	58.6	116	2	JC4669	cocaine-and amphet
136	34	58.6	131	2	F70422	hypothetical prote
137	34	58.6	142	2	T28467	hypothetical prote
138	34	58.6	151	2	C90275	hypothetical prote
139	34	58.6	155	2	G69877	hypothetical prote
140	34	58.6	166	2	A11444	hypothetical prote
141	34	58.6	190	2	G01808	HP1hs-alpha - huma
142	34	58.6	194	2	JC4951	rab protein 22B -
143	34	58.6	198	2	S68839	guanylate cyclase-
144	34	58.6	203	2	G86785	acetyltransferase
145	34	58.6	204	2	A57604	guanylate cyclase-
146	34	58.6	209	2	F82483	hypothetical prote
147	34	58.6	211	2	F69391	hypothetical prote
148	34	58.6	223	2	D84696	probable glucathio
149	34	58.6	228	2	C26599	clathrin light cha
150	34	58.6	229	2	B31775	clathrin light cha
151	34	58.6	229	2	S55963	hypothetical prote
152	34	58.6	229	2	C69531	competence-damage
153	34	58.6	238	1	LRR1B2	clathrin light cha
154	34	58.6	246	2	F97076	hypothetical prote
155	34	58.6	250	2	T20479	hypothetical prote
156	34	58.6	258	2	B82287	conserved hypotet
157	34	58.6	263	2	A30227	hypothetical prote
158	34	58.6	283	2	T31147	hypothetical prote
159	34	58.6	297	2	A98067	hypothetical prote
160	34	58.6	301	2	B72256	hypothetical prote
161	34	58.6	323	2	C82234	conserved hypotet
162	34	58.6	324	2	F71536	probable accoa car
163	34	58.6	324	2	D81691	acetyl-coenzyme A
164	34	58.6	327	2	A44013	gonadotropin-relea
165	34	58.6	338	2	S75196	hypothetical prote
166	34	58.6	350	2	H69770	C4-dicarboxylate b
167	34	58.6	352	2	A63610	a probable transmem
168	34	58.6	371	2	A61348	hypothetical prote
169	34	58.6	376	2	T28935	hypothetical prote
170	34	58.6	382	2	T20236	hypothetical prote
171	34	58.6	382	2	T01162	hypothetical prote
172	34	58.6	385	2	S64340	hypothetical prote
173	34	58.6	393	2	S48288	probable phosphopr
174	34	58.6	396	2	H96675	hypothetical prote
175	34	58.6	401	2	D71003	probable transamin
176	34	58.6	416	1	JC4952	transcription init
177	34	58.6	427	2	T06191	probable endo-1,4-
178	34	58.6	427	2	T06195	histidinol dehydro
179	34	58.6	429	2	B81802	histidinol dehydro
180	34	58.6	429	2	E81067	hypothetical prote
181	34	58.6	431	2	AH2006	hypothetical prote
182	34	58.6	442	2	H72209	hypothetical prote
183	34	58.6	446	1	C71011	hypothetical prote
184	34	58.6	446	2	D75120	hypothetical prote
185	34	58.6	449	2	G81419	hypothetical prote
186	34	58.6	459	2	B84933	UDP-N-acetylglucos
187	34	58.6	460	2	H83264	antithrombin III -
188	34	58.6	465	2	I59611	RNTI protein - yea
189	34	58.6	471	2	S56053	hypothetical prote
190	34	58.6	474	2	T47699	hypothetical prote
191	34	58.6	478	2	A66204	probable replicati
192	34	58.6	486	2	E71658	hypothetical prote
193	34	58.6	488	2	B83884	beta-xylosidase /
194	34	58.6	499	2	D97000	methyl-accepting c
195	34	58.6	517	2	I40798	cellulase (EC 3.2.
196	34	58.6	540	2	S76869	hypothetical prote
197	34	58.6	541	2	AP2483	hypothetical prote
198	34	58.6	549	1	S46474	glucan 1,4-alpha-g
199	34	58.6	550	2	I64203	aspartate-tRNA lig
200	34	58.6	564	2	S59649	ATP-dependent RNA
201	34	58.6	569	2	B83124	phytoene dehydroge
202	34	58.6	570	2	A39597	hypothetical prote
203	34	58.6	578	2	B71431	hypothetical prote
204	34	58.6	584	2	F84425	related to heterok
205	34	58.6	584	2	T49320	conserved hypotet
206	34	58.6	619	2	F70480	GTP-binding protel
207	34	58.6	645	2	S50374	NADPH-ferrithempro
208	34	58.6	683	2	A85044	protein F7F2.18 f
209	34	58.6	711	2	T14081	aspartokinase II/h
210	34	58.6	712	2	S21531	hypothetical prote
211	34	58.6	743	2	H96493	dynamin-like prote
212	34	58.6	803	2	B82045	DNA mismatch repai
213	34	58.6	860	2	E85074	hypothetical prote
214	34	58.6	920	2	T52426	hypothetical prote
215	34	58.6	934	2	I64819	hypothetical prote
216	34	58.6	946	2	T19465	hypothetical prote
217	34	58.6	985	2	T06049	protein F14N23.17
218	34	58.6	992	2	B86237	serine/threonine p
219	34	58.6	1001	2	T17365	hypothetical prote
220	34	58.6	1064	2	T01960	carbamoyl-phosphat
221	34	58.6	1089	2	E81446	protein T30E16.17
222	34	58.6	1092	2	H96619	DNA polymerase III
223	34	58.6	1161	2	B70387	probable swi/snf h
224	34	58.6	1163	2	E71481	swi/snf family hel
225	34	58.6	1166	2	G86596	hypothetical prote
226	34	58.6	1166	2	C72027	hypothetical prote
227	34	58.6	1187	2	JE0347	156k protein - Pla
228	34	58.6	1206	2	AC3140	mitotic spindle pr
229	34	58.6	1385	2	S34230	hypothetical prote
230	34	58.6	1428	2	C85079	hypothetical prote
231	34	58.6	1513	2	S45768	hypothetical prote
232	34	58.6	1547	2	JC0096	hypothetical prote
233	34	58.6	1896	2	T01490	hypothetical prote
234	34	58.6	1899	2	T32732	PAM C-terminal int
235	34	58.6	1919	2	T42098	PAM interacting pr
236	34	58.6	2091	2	A97077	hypothetical prote
237	34	58.6	2186	2	T31369	ligandin - fruit fl
238	34	58.6	2606	2	T03159	large tegument pro
239	34	58.6	2629	2	I46569	apolipoprotein B -
240	34	58.6	3071	2	T45584	apolipoprotein B -
241	34	58.6	4563	1	LPHUB	hypothetical prote
242	34	58.6	4578	1	S47817	apolipoprotein B-1
243	33.5	57.8	153	2	S47817	hypothetical prote
244	33.5	57.8	153	2	G91187	hypothetical prote
245	33.5	57.8	153	2	H86034	hypothetical prote
246	33	57.8	954	2	G71496	hypothetical prote
247	33	56.9	62	2	B69258	hypothetical prote
248	33	56.9	96	2	H87336	conserved hypotet
249	33	56.9	100	2	G97848	hypothetical prote

249	33	56.9	127	2	T10546	Rab1 protein homo
250	33	56.9	136	2	A11104	B. subtilis yacC p
251	33	56.9	136	2	A11466	B. subtilis yacC p
252	33	56.9	138	2	AD1850	hypothetical prote
253	33	56.9	142	2	T45012	hypothetical prote
254	33	56.9	143	2	E71041	hypothetical prote
255	33	56.9	152	2	D64839	probable phosphat
256	33	56.9	152	2	B90771	probable phosphat
257	33	56.9	152	2	F85633	probable phosphat
258	33	56.9	154	2	D42148	GTP-binding protei
259	33	56.9	155	2	A70326	conserved hypochet
260	33	56.9	155	2	T14781	hypothetical prote
261	33	56.9	172	2	C83269	hypothetical prote
262	33	56.9	194	2	C70209	conserved hypochet
263	33	56.9	198	2	AC1890	hypothetical prote
264	33	56.9	202	2	A61379	22k antigen - Rick
265	33	56.9	203	2	A46447	GTP-binding protei
266	33	56.9	203	2	F90658	hypothetical prote
267	33	56.9	203	2	T27592	hypothetical prote
268	33	56.9	203	2	F85509	hypothetical prote
269	33	56.9	209	2	B90497	conserved hypochet
270	33	56.9	210	2	AC2402	KHG-KDPG bifunctio
271	33	56.9	210	2	G84899	hypothetical prote
272	33	56.9	214	2	D83278	hypothetical prote
273	33	56.9	214	2	A12371	hypothetical prote
274	33	56.9	224	2	F84696	probable glutathio
275	33	56.9	225	2	G84696	probable glutathio
276	33	56.9	232	2	S40165	glutathione transf
277	33	56.9	242	2	S70176	yeast protein - Yer
278	33	56.9	245	2	AB0300	quorum-sensing tra
279	33	56.9	246	2	C97177	C-terminal domain
280	33	56.9	258	2	A64318	hypothetical prote
281	33	56.9	259	2	E69863	conserved hypochet
282	33	56.9	259	2	B90535	hypothetical prote
283	33	56.9	266	2	H90286	hypothetical prote
284	33	56.9	266	2	S48409	hypothetical prote
285	33	56.9	269	2	F69541	spermidine synthas
286	33	56.9	277	2	B82838	conserved hypochet
287	33	56.9	277	2	T27597	hypothetical prote
288	33	56.9	282	2	F72595	hypothetical prote
289	33	56.9	283	2	A71228	probable agmatinas
290	33	56.9	284	2	F97216	probable aminoglyc
291	33	56.9	297	2	T12615	ribosomal protein
292	33	56.9	300	2	D83714	surface adhesin A
293	33	56.9	302	2	A72510	hypothetical prote
294	33	56.9	309	2	E69200	conserved hypochet
295	33	56.9	321	2	G90991	probable kinase [1
296	33	56.9	321	2	A85837	probable kinase [1
297	33	56.9	322	2	E71137	hypothetical prote
298	33	56.9	322	2	G64388	hypothetical prote
299	33	56.9	325	2	T24834	hypothetical prote
300	33	56.9	327	2	I60169	gonadotropin-relea

ALIGNMENTS

RESULT 1
G70478
hypothetical protein aq_2086 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: G70478
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: G70478
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1118 <NOF>
A:Cross-references: GB:AE000772; NID:92984299; PIDN:AAC07832.1; PID:92984309; GB:AE00065

A:Experimental source: strain VFS
C:Genetics:
A:Gene: aq_2086
C:Superfamily: conserved hypothetical protein MJ1624

Query Match 72.4%; Score 42; DB 2; Length 118;
Best Local Similarity 70.0%; Pred. No. 8.5;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 LEWKLKKR 10
| | | | |
DB 8 LSEWKLKKR 17

RESULT 2

7,8-diaminopelargonic acid synthetase [imported] - Escherichia coli (strain O157:H7,
E85585
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 30-Sep-2001
C:Accession: E85585
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Diallanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: E85585

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-429 <STO>

A:Cross-references: GB:AE005174; NID:912513760; PIDN:AAG55145.1; GSPDB:GN00145; UWGP:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:

A:Gene: hlyA
C:Superfamily: ornithine-oxo-acid aminotransferase

Query Match 70.7%; Score 41; DB 2; Length 429;
Best Local Similarity 66.7%; Pred. No. 46;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 EWLKRLKKM 11
| | | | |
DB 226 EWLKRLKKM 234

RESULT 3

7,8-diaminopelargonic acid synthetase [imported] - Escherichia coli (strain O157:H7,
D90735
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 17-May-2002
C:Accession: D90735
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurikawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhtara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and 9
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: D90735
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-429 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA834275.1; PID:913360311; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: E850852
C:Superfamily: ornithine-oxo-acid aminotransferase

Query Match 70.7%; Score 41; DB 2; Length 429;
Best Local Similarity 66.7%; Pred. No. 46;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 EWLKRLKKM 11
| | | | |
DB 226 EWLKRLKKM 234

Fri Jul 11 16:11:41 2003

us-09-828-592-10.rpr

RESULT 4

AF3270 Integrase [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C:Spec: 01-Feb-2002 #sequence-revision 01-Feb-2002 #text-change 15-Feb-2002

C:Date: 01-Feb-2002 #sequence-revision 01-Feb-2002 #text-change 15-Feb-2002

C:Accession: AF3270

R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AD3252; PMID:11756688

A:Accession: AF3270

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-315 <KUR>

A:Cross-references: GB:AE008917; PIDN:AU51329.1; PID:g17982026; GSPDB:GN00190

A:Experimental source: strain 16M

A:Gene: BM10147

C:Superfamily: probable site-specific integrase/recombinase Xerc

A:Map position: 1

Query Match 69.0%; Score 40; DB 2; Length 315;

Best Local Similarity 70.0%; Pred. No. 48;

Matches 7; Conservative 1; Mismatches 0; Gaps 0;

Db 2 OEMLKLLKKM 11

Db 20 EEWKSLKTM 29

RESULT 5

F71172 Hypothetical protein PH0578 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C:Spec: 14-Aug-1998 #sequence-revision 14-Aug-1998 #text-change 21-Jul-2000

C:Date: 14-Aug-1998 #sequence-revision 14-Aug-1998 #text-change 21-Jul-2000

C:Accession: F71172

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Hatakeyama, Y.; Hino, Y.; Yamamoto, S.; Sekin,

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kusuda, N.; Oguchi,

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A:Reference number: A71000; MUID:98344137; PMID:9679194

A:Accession: F71172

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-457 <KAW>

A:Cross-references: GB:AP000002; NID:93236129; PIDN:BAA29667.1; PID:93256984

A:Experimental source: strain 073

A:Note: this accession replaces an interim accession for a sequence replaced by Genbank

C:Gene: PH0578

Query Match 69.0%; Score 40; DB 2; Length 457;

Best Local Similarity 54.5%; Pred. No. 70;

Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 1 OEMLKLLKKM 11

Db 445 LEFWKCKEKL 455

RESULT 6

XHHU3 Anthrombin III precursor [validated] - human

C:Species: Homo sapiens (man)

C:Spec: 05-Apr-1983 #sequence-revision 05-Apr-1983 #text-change 08-Dec-2000

C:Date: 05-Apr-1983 #sequence-revision 05-Apr-1983 #text-change 08-Dec-2000

C:Accession: A49494; A93453; A92431; A94445; A94190; S02530; S63600; I37

R:Olds, R.J.; Lane, D.A.; Chowdhury, V.; De Stefano, V.; Leone, G.; Rhet, S.L.

Biochemistry 37, 4216-4224, 1993

A:Title: Complete nucleotide sequence of the antithrombin gene: evidence for homologous

A:Reference number: A49494; MUID:93237227; PMID:8476848

A:Accession: A49494

A:Molecule type: DNA

A:Residues: 1-464 <OLD>

A:Cross-references: EMBL:X68793; NID:928906; PIDN:CAA48690.1; PID:928907

A:Note: sequence extracted from NCB1 backbone (NCBIN:130247, NCBIF:130246)

A:Rock, S.C.; Wilson, K.L.; Vohar, G.A.; Lavin, R.M.

Nucleic Acids Res. 10, 8113-8125, 1982

A:Title: Cloning and expression of the cDNA for human antithrombin III.

A:Reference number: A93453; MUID:83143280; PMID:6298709

A:Accession: A93453

A:Molecule type: mRNA

A:Residues: 1-464 <ROCI>

A:Cross-references: GB:J00190; GB:J00103; GB:J00104; NID:9179128; PIDN:AAB

R:Chandra, T.; Stackhouse, R.; Kidd, V.J.; Moo, S.L.C.

Proc. Natl. Acad. Sci. U.S.A. 80, 1845-1848, 1983

A:Title: Isolation and sequence characterization of a cDNA clone of human antithrombi

A:Reference number: A93943; MUID:83169777; PMID:6572945

A:Accession: A93943

A:Molecule type: mRNA

A:Residues: 1-464 <CHRA>

A:Cross-references: GB:J00190; GB:J00102; GB:J00103; GB:J00104; NID:9179128; PIDN:AAB

R:Prochownik, E.V.; Markham, A.F.; Orkin, S.H.

J. Biol. Chem. 258, 8389-8394, 1983

A:Title: Isolation of a cDNA clone for human antithrombin III.

A:Reference number: A92431; MUID:83238456; PMID:6305982

A:Accession: A92431

A:Molecule type: mRNA

A:Residues: 42-96, 'R', '98-464 <PRO>

A:Note: the authors translated the codon GAC for residue 206 as Asn, GGC for residue

R:Chang, J.Y.

J. Biol. Chem. 264, 3111-3115, 1989

A:Title: Binding of heparin to human antithrombin III activates selective chemical mo

in III.

A:Reference number: A33305; MUID:89123426; PMID:2492530

A:Accession: A33305

A:Molecule type: Protein

A:Residues: 90-105,124-143,147-161,165-171,268-273,381-402,446-457 <CH2>

A:Reference number: A94445

A:Accession: A94445

A:Molecule type: Protein

A:Residues: 33-66, 'OE', '71-242,243-245,248-250,254-464 <DET>

A:Note: carbonyl-ester-binding site and disulfide bonds

R:Zetlin, G.; Contrad, H.S.; Nimitz, M.; Kargue, H.E.

J. Biol. Chem. 264, 21153-21159, 1989

A:Title: Characterization of recombinant human antithrombin III synthesized in Chines

A:Reference number: A34190; MUID:90078215; PMID:2592368

A:Accession: A34190

A:Molecule type: Protein

A:Residues: 33-39 <ZET>

A:Experimental source: recombinant protein from Chinese hamster ovary (CHO) cells

A:Reference number: S02530; MUID:87304255; PMID:3305015

A:Accession: S02530

A:Status: preliminary

A:Molecule type: Protein

A:Residues: 33-39, 'X', '41-42,67-77, 'XX', '80-88, 'X', '90-92,137-142,189-202,210-214, 'X', '26

R:Chang, W.S.W.; Wardell, M.R.; Lomas, D.A.; Carrell, R.W.

Biochem. J. 314, 647-653, 1996

A:Title: Probing the heparin-binding domain of human antithrombin III with V8 proteas

A:Reference number: S63599; MUID:96239126; PMID:8670081

A:Accession: S63599

A:Status: preliminary

A:Molecule type: Protein

A:Residues: 414-428 <CH3>

R:Rock, S.C.; Levitan, D.J.

Nucleic Acids Res. 11, 8569-8582, 1983

A:Title: Characterization of an unusual DNA length polymorphism 5' to the human antit

A:Reference number: I37191; MUID:84169500; PMID:6672771

A:Accession: I37191

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-14 <BOC2>
A:Cross-references: EMBL:X00237; NID:g28917; PIDN:CAA5059.1; PID:g28918; EMBL:X00238; N
Proc. Natl. Acad. Sci. U.S.A. 81, 289-293, 1984
A:Title: Antithrombin III Toyama: replacement of arginine-47 by cysteine in hereditary
A:Reference number: A29371; MUID:84119472; PMID:6582486
A:Accession: A29371
A:Molecule type: protein
A:Residues: 53-68, 'OE', '71-78', 'C', '80-135' <K01>
A:Note: variant Toyama with a substitution of Cys for Arg-79 and consequently cannot b
R:Boek, S.C.; Marrian, J.A.; Radziejewska, E.
Biochemistry 27, 6171-6178, 1988
A:Title: Antithrombin III Utah: proline-407 to leucine mutation in a highly conserved re
A:Reference number: I52399; MUID:89050967; PMID:3191114
A:Note: a published extractum appears in Biochemistry 28, 3628 1989
A:Accession: I52399
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 137-208 <BOC3>
A:Cross-references: GB:M21643; NID:g179147; PIDN:AAA51793.1; PID:g457132
A:Accession: I65277
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 137-254 <BOC4>
A:Cross-references: GB:M21644; NID:g179148; PIDN:AAA51794.1; PID:g179151
A:Accession: I65278
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 407-464 <BOC5>
A:Cross-references: GB:M21645; NID:g179149; PIDN:AAA51795.1; PID:g179152
A:Accession: I65279
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-438, 'L', '440-464' <BOC6>
A:Cross-references: GB:M21642; NID:g179159; PIDN:AAA51796.1; PID:g179161
A:Note: mutant Utah
R:Boek, S.C.; Brennan, S.O.; Carrell, R.W.; George, P.; Perry, D.J.; Shaw, J.
FEBS Lett. 266, 163-166, 1990
A:Title: Antithrombin rosen-IV 24 Arg->Cys. The amino-terminal contribution to heparin b
A:Reference number: S10716; MUID:90306344; PMID:2365065
A:Accession: S10716
A:Molecule type: protein
A:Residues: 53-61 <BOB>
A:Note: variant form Rouen-IV, 56-Cys, was also sequenced
R:Gundry, C.B.; Thomas, F.; Miller, D.S.; Krawczak, M.; Mellisart, E.; Lindo, V.; Moffat
Blood 78, 1027-1032, 1991
A:Title: Recurrent deletion in the human antithrombin III gene.
A:Reference number: A44935; MUID:91329813; PMID:1868237
A:Accession: A44935
A:Molecule type: DNA
A:Residues: 271-276, 'GRVQHT', <GR1>
A:Cross-references: GB:S49757; NID:g233568; PIDN:AA19467.1; PID:g233569
A:Experimental source: antithrombin III deficiency patient 1
A:Note: sequence extracted from NCBI backbone (NCBIN:49757, NCBIIP:49758)
A:Accession: B44935
A:Molecule type: DNA
A:Residues: 271-276, 'VVSIVDVRQVPLSARG', <GR2>
A:Cross-references: GB:S49759; NID:g233570; PIDN:AA19468.1; PID:g233571
A:Experimental source: antithrombin III deficiency patient 2
A:Note: sequence extracted from NCBI backbone (NCBIN:49759, NCBIIP:49760)
A:Note: different frameshift mutations at an identical site in unrelated patients sugges
Thromb. Haemost. 67, 521-525, 1992
A:Title: Insertions/deletions in the antithrombin gene: 3 mutations associated with non-
A:Reference number: I59610; MUID:92390894; PMID:1325679
A:Accession: I59610
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 79, 'CLGTVQGFPLCYHLSAPGRQF', <DAL1>
A:Cross-references: GB:S43612; NID:g234806; PIDN:AA23132.1; PID:g234807
A:Note: frameshift mutant, patient 1
A:Accession: I81229
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 401, 'RHPLR', <DAL2>
A:Cross-references: GB:S43621; NID:g254810; PIDN:AA23134.1; PID:g254811
A:Note: frameshift mutant, patient 3
R:Boek, S.C.; Danielsson, A.; Fenton II, J.W.; Jorvall, H.
FEBS Lett. 126, 257-260, 1981
A:Title: The site in human antithrombin for functional proteolytic cleavage by human
A:Reference number: A91287; MUID:8121814; PMID:723875
A:Contents: annotation: inhibitory site
R:Blackburn, M.N.; Smith, R.L.; Carson, J.; Sibley, C.C.
J. Biol. Chem. 259, 939-941, 1984
A:Title: The heparin-binding site of antithrombin III. Identification of a critical t
A:Reference number: A92488; MUID:8411578; PMID:663405
A:Contents: annotation: heparin-binding site
C:Genetics:
A:Gene: GDB:AT3
A:Cross-references: GDB:119024; OMIM:107300
A:Map position: 1923-1925.1
A:Introns: 14/2; 136/3; 208/3; 254/3; 385/1; 406/3
C:Function:
A:Description: in blood plasma inhibits thrombin and activated coagulation factor X,
C:Superfamily: antithrombin III
C:Keywords: acute phase; glycoprotein; heparin binding; plasma; serine proteinase inh
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-464/Product: antithrombin III #status experimental <MAT>
F:40-160,53-127,279-462/Disulfide bonds: #status experimental
F:81/Binding site: heparin (TRP) #status experimental
F:128,167,187,224/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:445-426/Cleavage site: Arg-Ser (thrombin) #status experimental
F:425/inhibitory site: Arg (thrombin, coagulation factor Xa) #status experimental
Query Match 69.0%; Score 40; DB 1; Length 464;
Best Local Similarity 63.6%; Pred. No. 71;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Caps 0;
OY 1 LOEWLKKKKM 11
DB 336 LOEWLDELEM 346
RESULT 7
57851
probable aspartate-tRNA ligase (EC 6.1.1.12) - Mycoplasma capricolum
N:Alternate names: aspartyl-tRNA synthetase; protein MC054; protein MC191
C:Species: Mycoplasma capricolum
C:Date: 09-Oct-1997 #sequence-revision 15-May-1998 #text-change 03-Jun-2002
A:Accession: S77851; S77845
R:Boek, S.C.; Ouzounis, C.; Casati, G.; Schneider, R.; Sander, C.; Dolan, M.; Gilbert,
Mol. Microbiol. 16, 955-967, 1995
A:Title: Exploring the Mycoplasma capricolum genome: a minimal cell reveals its phys
A:Reference number: S77739; MUID:96059641; PMID:7476192
A:Accession: S77851
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-51 <BOB>
A:Cross-references: EMBL:Z33137; NID:g516219; PIDN:CAA83775.1; PID:g530448
A:Experimental source: ATCC 27343
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
A:Accession: S77845
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 42-524 <BOB>
A:Cross-references: EMBL:Z33048; NID:g514447; PIDN:CAA83720.1; PID:g530412
A:Experimental source: ATCC 27343
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Genetics:
A:Genetic code: SGC3
C:Superfamily: lysine-tRNA ligase
C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis
Query Match 69.0%; Score 40; DB 2; Length 524;

Best Local Similarity 54.5%; Pred. No. 81;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LOEWLKKLKKM 11
||:||||:
Db 21 LQWKVKIRKL 31

RESULT 8

H85048

probable transposon protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001

C:Accession: H85048

R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring

Nature 402: 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; MUID:20083488; PMID:10617198

A:Accession: H85048

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-745 <STO>

A:Cross-references: GB:NC_001268; NID:g7267141; PIDN:CAB80809.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4g03860

A:Map position: 4

Query Match 69.0%; Score 40; DB 2; Length 745;
Best Local Similarity 63.6%; Pred. No. 1.2e+02;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LOEWLKKLKKM 11
||:||||:
Db 593 LQWKVKIRKL 603

RESULT 9

T01863

hypothetical protein T7M24.3 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 24-Mar-1999

C:Accession: T01863

R:Harmon, G.; Langston, Y.; Stoneking, T.; Drone, K.; Ames, M.

submitted to the EMBL Data Library, July 1998

A:Description: The sequence of Arabidopsis thaliana T7M24.

A:Reference number: Z14448

A:Accession: T01863

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1120 <HAR>

A:Cross-references: EMBL:AF077408; NID:g3319359; PID:g3319360

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 4

A:Introns: 76/2; 118/3; 421/2; 555/3; 623/1; 752/3; 908/3

A:Note: T7M24.3

Query Match 69.0%; Score 40; DB 2; Length 1120;
Best Local Similarity 63.6%; Pred. No. 1.7e+02;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LOEWLKKLKKM 11
||:||||:
Db 968 LQWKVKIRKL 978

RESULT 10

A82939

membrane nuclease A, hypothetical U0055 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: A82939

R:Glass, J.I.; Leftkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.

submitted to Genbank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a

A:Reference number: A82870

A:Accession: A82939

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1138 <GLA>

A:Cross-references: GB:AE002105; GB:AF222894; NID:g6899003; PIDN:AAF30460.1; GSPDB:GN

C:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: mnuA-1; U0055

A:Genetic code: SGC3

Query Match 69.0%; Score 40; DB 2; Length 1138;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 QEWLKKLKKM 11
|||||:
Db 1095 QEWLKKVRL 1104

RESULT 11

RRW6SM

RNA-directed RNA polymerase (EC 2.7.7.48) - strawberry mild yellow edge-associated vi

N:Alternate names: RNA nucleotidyltransferase (RNA-directed); RNA replicase

C:Species: strawberry mild yellow edge-associated virus

C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 19-Jan-2001

C:Accession: J01426

R:Jelkmann, W.; Maiss, E.; Martin, R.R.

J. Gen. Virol. 73, 475-479, 1992

A:Title: The nucleotide sequence and genome organization of strawberry mild yellow ed

A:Reference number: J01426; MUID:92166762; PMID:1339469

A:Accession: J01426

A:Molecule type: genomic RNA

A:Residues: 1-1323

A:Cross-references: GB:D12517; DBJ:D01227; NID:g222631; PIDN:BAA02082.1; PID:g222632

C:Superfamily: eggplant mosaic virus RNA-directed RNA polymerase

C:Keywords: ATP; nucleotide binding; nucleotidyltransferase; P-loop; RNA biosynthesis

F:605-612/Region: nucleotide-binding motif A (P-loop)

F:668-673/Region: nucleotide-binding motif B

F:611/Binding site: ATP (Lys) #status predicted

Query Match 69.0%; Score 40; DB 1; Length 1323;
Best Local Similarity 60.0%; Pred. No. 2.1e+02;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LOEWLKKLKK 10
||:||||:
Db 615 LODMRRSLKK 624

RESULT 12

H85062

hypothetical protein AT4g05000 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001

C:Accession: H85062

R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp

Nature 402: 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; MUID:20083488; PMID:10617198

A:Accession: H85062

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-209 <STO>

A:Cross-references: GB:NC_001268; NID:g7267259; PIDN:CAB81042.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4g05000

A:Map position: 4

Query Match 67.2%; Score 39; DB 2; Length 209;
Best Local Similarity 54.5%; Pred. No. 46;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LQEWLKKLKK 11
::|||::|||
Db 165 MKEWLSRLSKM 175

RESULT 13

T46279
hypothetical protein DKFZP56401278.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C:Accession: T46279

R:Ottensmøder, B.; Obermayer, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A:Reference number: 223031

A:Accession: T46279

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-314 <AAA>

A:Cross-references: EMBL:AL137517

A:Experimental source: fetal brain; clone DKFZP56401278

C:Genetics:

A:Note: DKFZP56401278.1

Query Match 67.2%; Score 39; DB 2; Length 314;
Best Local Similarity 60.0%; Pred. No. 69;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LQEWLKKLKK 10
|||::|||
Db 1 LQWIKRLSK 10

RESULT 14

H83977
amino transferase BH2624 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: H83977

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: H83977

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-454 <STO>

A:Cross-references: GB:BA000004; NID:g10175192; PIDN:BAH06343.1; GSPDB:GNOC

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH2624

C:Superfamily: ornithine-oxo-acid aminotransferase

Query Match 67.2%; Score 39; DB 2; Length 454;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 EWLKKLKK 11
|||::|||
Db 238 EWLKELRRM 246

RESULT 15

D75053
hypothetical protein PAB0938 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: D75053

R:Anonymous; Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A:Reference number: A75001

A:Accession: D75053
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-480 <KAM>
A:Cross-references: GB:AJ248287; GB:AL096836; NID:g5458657; PIDN:CAH50321.1; PID:g545
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB0938

Query Match 67.2%; Score 39; DB 2; Length 480;
Best Local Similarity 54.5%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LQEWLKKLKK 11
|||::|||
Db 468 LQWVKCKKL 478

RESULT 16

T14776
hypothetical protein DKFZP434N101.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000

C:Accession: T14776

R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, August 1999

A:Reference number: 218183

A:Accession: T14776

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-629 <POU>

A:Cross-references: EMBL:AL110247

A:Experimental source: adult testis; clone DKFZP434N101

C:Genetics:

A:Note: DKFZP434N101.1

C:Superfamily: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase II; 1-phospho

hodiesterase domain Y homology; SH2 homology; SH3 homology

Query Match 67.2%; Score 39; DB 2; Length 629;
Best Local Similarity 45.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1 LQEWLKKLKK 11
|||::|||
Db 204 LQWVKKIREV 214

RESULT 17

F64526
adenine/cytosine DNA methyltransferase - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999

C:Accession: F64526

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.

Peterson, S.; Loftus, B.; Richardson, D.; Khakhria, H.G.; Glodok, A.; McKe

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney,

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser,

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: F64526

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-823 <TOM>

A:Cross-references: GB:AE000527; GB:AE000511; NID:g2313131; PIDN:AAD07124.1; PID:g231

Query Match 67.2%; Score 39; DB 2; Length 823;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LQEWLKKLKK 10
|||::|||
Db 246 MQEWIKLKE 255

RESULT 18

AH0950

DNA polymerase I [imported] - Salmonella enterica subsp. enterica serovar Typh1 (strain C:Species: Salmonella enterica subsp. enterica serovar Typh1)

A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001

C:Accession: AH0950

R:Parkhill, J., Dougan, G., James, K.D., Thomson, N.R., Pickard, D., Wain, J., Churcher, Th., T., Conerton, P., Cronin, A., Davis, P., Davies, R.M., Dowd, L., White, N., Farrar, S., Moule, S., O'Garra, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C., Quail, M., Rutherford, K., Simmonds, M., Skelton, J., Stevens, K., A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; PMID:11677608

A:Accession: AH0950

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-928 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD03100.1; PID:916504737; GSPDB:GN00176

C:Genetics:

A:Gene: STY3881

C:Superfamily: DNA-directed DNA polymerase I

Query Match 67.2%; Score 39; DB 2; Length 928;

Best Local Similarity 60.0%; Pred. No. 2.1e+02;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LQEWLKLKK 10

Db 339 LEWIKLKK 348

RESULT 19

E84022

hypothetical protein BH2981 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: E84022

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: E84022

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-955 <STO>

A:Cross-references: GB:AP001517; GB:BA000004; NID:910175500; PIDN:BA06700.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH2981

Query Match 67.2%; Score 39; DB 2; Length 955;

Best Local Similarity 66.7%; Pred. No. 2.2e+02;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 QEWLKLKK 10

Db 748 REWLKLKK 756

RESULT 20

B86423

hypothetical protein TIP2.15 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001

C:Accession: B86423

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpsiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Hultzai, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Matli, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salberg, S.L.; Schartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: B86423

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-1114 <STO>

A:Cross-references: GB:AE005172; NID:910092454; PIDN:AA012856.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 67.2%; Score 39; DB 2; Length 1114;

Best Local Similarity 54.5%; Pred. No. 2.5e+02;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LQEWLKLKK 11

Db 328 LEWLKCKSL 338

RESULT 21

A31317

1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase (EC 3.1.4.11) II - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 31-Mar-2000

C:Accession: A31317; A56761

R:Sun, P.G.; Ryu, S.H.; Moon, K.H.; Suh, H.W.; Rhee, S.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 5419-5423, 1988

A:Title: Inositol phospholipid-specific phospholipase C: complete cDNA and protein se

A:Reference number: A31317; MUID:88289733; PMID:2840660

A:Accession: A31317

A:Molecule type: mRNA

A:Residues: 1-1290 <SUH>

A:Cross-references: GB:J03806; NID:9206323; PIDN:AAA41921.1; PID:9206324

A:Rhee, S.J.; Ryu, S.H.; Suh, P.G.

Biochem. Biophys. Res. Commun. 194, 294-300, 1993

A:Title: Promoter region of the rat phospholipase C-gamma 1 gene.

A:Reference number: A56761; MUID:9326133; PMID:8392838

A:Accession: A56761

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-29 <LEP>

A:Note: Sequence extracted from NCBI backbone (NCBIN:135515, NCBI:135516)

C:Superfamily: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase II; 1-phosph

hodiesterase domain y homology; SH2 homology; SH3 homology

C:Keywords: phosphoprotein; phosphoric diester hydrolase

F:322-464/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain x

F:550-657/Domain: SH2 homology <SH2>

F:668-756/Domain: SH2 homology <SH2>

F:798-846/Domain: SH3 homology <SH3>

F:952-1073/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain y

Query Match 67.2%; Score 39; DB 2; Length 1290;

Best Local Similarity 45.5%; Pred. No. 2.9e+02;

Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1 LQEWLKLKK 11

Db 919 LQDWKKIREV 929

RESULT 22

A36466

1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase (EC 3.1.4.11) - human

C:Species: Homo sapiens (man)

C:Date: 15-Feb-1991 #sequence_revision 15-Feb-1991 #text_change 04-Feb-2000

C:Accession: A36466

R:Burgess, W.H.; Dione, C.A.; Kaplow, J.; Mudd, R.; Friesel, R.; Zilberstein, A.; Sc

Mol. Cell. Biol. 10, 4770-4777, 1990

```

A>Title: Characterization and cDNA cloning of phospholipase C-gamma, a major substrate
A:Reference number: A36466; MUID:90355993; PMID:2167438
A:Accession: A36466
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1290 <BUR>
A:Cross-references: GB:M4667; NID:g190037; PTDN:AAA36452.1; PID:g190038
C:Genetics:
A:Gene: GDB:PLCG1; PLC1
A:Cross-references: GDB:120299; OMIM:172420
C:Map position: 20q12-20q13.1
A:Superfamily: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase II; 1-phosphatidylcholine-specific phospholipase C
C:Keywords: domain y homology; SH2 homology; SH3 homology
F:322-464/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain x homolog
F:550-657/Domain: SH2 homology <SH21>
F:668-756/Domain: SH2 homology <SH22>
F:798-846/Domain: SH3 homology <SH3>
F:952-1073/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain y homolog
Query Match          67.2%; Score 39; DB 2; Length 1290;
Best Local Similarity 45.5%; Pred. No. 2.9e+02;
Matches      5; Conservative    6; Mismatches     0; Indels       0; Gaps        0;

OY      1 LQEWLKKLKRM 11
|||:|||:::
Db      919 LQDWVKKIREV 929

RESULT 23
S00666
1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase (EC 3.1.4.11) Isozyme II - bovin
N:Alternate names: phosphoinositol-specific phospholipase C-II; phospholipase C, 148K
C:Species: Bos primigenius taurus (cattle)
C>Date: 31-Dec-1988 #sequence_revison 31-Dec-1988 #text_change 04-Feb-2000
C:Accession: S00666
R:Stahl,M.L.; Ferez,C.R.; Kelleher,K.L.; Kriz,R.W.; Knopf,J.L.
Nature 332, 269-272, 1988
A>Title: Sequence similarity of phospholipase C with the non-catalytic region of src.
A:Reference number: S00666; MUID:88156963; PMID:2831461
A:Accession: S00666
A:Molecule type: mRNA
A:Residues: 1-1291 <STRA>
A:Cross-references: EMBL:X00301; NID:9638; PTDN:CAA68406.1; PID:9639
C:Superfamily: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase II; 1-phosphatidylcholine-specific phospholipase C
C:Keywords: domain y homology; SH2 homology; SH3 homology
F:322-464/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain x homolog
F:550-657/Domain: SH2 homology <SH21>
F:668-756/Domain: SH2 homology <SH22>
F:798-846/Domain: SH3 homology <SH3>
F:952-1073/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain y homolog
Query Match          67.2%; Score 39; DB 2; Length 1291;
Best Local Similarity 45.5%; Pred. No. 2.9e+02;
Matches      5; Conservative    6; Mismatches     0; Indels       0; Gaps        0;

OY      1 LQEWLKKLKRM 11
|||:|||:::
Db      919 LQDWVKKIREV 929

RESULT 24
D86798
Prophage p13 protein 14 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: D86798
C:Gene:
R:Boloitin, A.; Wincker, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich,
Genome Res. 11, 731-753, 2001
A>Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis sensu stricto
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: D86798

```

```

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1640 <STO>
A:Cross-references: GB:AE005176; PID:912724374; PIDN:AAK05486.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: plj14

Query Match      67.2%, Score 39; DB 2; Length 1640;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 WLKKLKK 10
      |||||
DB      671 WLKKLKK 677

RESULT 25
OPCH2
acylphosphatase (EC 3.6.1.7) Ch2, skeletal muscle - chicken
N:Alternate names: acylphosphate phosphohydrolase
C:Species: Gallus gallus (chicken)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Aug-1994
C:Accession: A41513
R:Ohba, Y.; Minowa, O.; Mizuno, Y.; Shlokawa, H.
J. Biochem. 102, 1221-1229, 1987
A:Title: The primary structure of chicken muscle acylphosphatase isozyme Ch2.
A:Reference number: A41513; MUID:88139274; PMID:2830254
A:Accession: A41513
A:Molecule type: protein
A:Residues: 1-98 <OH>
C:Superfamily: acylphosphatase
C:Keywords: acetylated amino end; hydrolase
F.1/Modified site: acetylated amino end (Ala) #status experimental

Query Match      65.5%, Score 38; DB 1; Length 98;
Best Local Similarity 75.0%; Pred. No. 31;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 LOEWLKKL 8
      |||||
DB      61 LOEWLKKL 68

RESULT 26
G69442
transcription regulator Atsr family homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C:Accession: G69442
R:Klink, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.; Glodek, A.; Zhou, L.; Overbeek, R.; Goeyne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Attlich, P.; Kaine, B.P.; Sykes, Smith, H.O.; Weese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
A:Accession number: A69250; MUID:98049343; PMID:9389475
A:Accession: G69442
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-131 <KLE>
A:Cross-references: GB:AE000996; GB:AE000782; NID:92689319; PIDN:AA889703.1; PID:926646

Query Match      65.5%, Score 38; DB 2; Length 131;
Best Local Similarity 66.7%; Pred. No. 41;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      2 QEWLKKLKK 10
      :|||:
DB      32 KEWLKKLR 40

```

RESULT 27
C75173
hypothetical protein PAB2001 - *Pyrococcus abyssi* (strain Orsay)
A:Species: *Pyrococcus abyssi*
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: C75173
R:anonymous, Genoscope
Submitted to the EMBL Data Library, July 1999
A:Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome structure
A:Reference number: A75001
A:Accession: C75173
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-143 <KAM>
A:Cross-references: GB:AJ248284; GB:AL096836; NID:g5457730; PIDN:CAB49466.1; PID:e151536
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB2001

Query Match 65.5%; Score 38; DB 2; Length 143;
Best Local Similarity 60.0%; Pred. No. 45;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOEWLKLKK 10
|:|||||:
DB 113 LKWLKRYE 122

RESULT 28
A64492
toxin sensitivity protein KTI12 homolog - *Methanococcus jannaschii*
C:Species: *Methanococcus jannaschii*
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: A64492
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reisch, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Ison, J.D.; Sadov, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: A64492
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-252 <BOU>
A:Cross-references: GB:U67594; GB:L77117; NID:g1592160; PIDN:AAB99557.1; PID:g1592169; T
C:Genetics:
A:Map position: FOR1515893-1516551
A:Start codon: TTG

Query Match 65.5%; Score 38; DB 2; Length 252;
Best Local Similarity 70.0%; Pred. No. 80;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 OEWLKLLKK 11
:|||||:
DB 224 KEFLKIRKM 233

RESULT 29
S61990
hypothetical protein YOR006c - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: hypothetical protein O2533; hypothetical protein UND313
C:Species: *Saccharomyces cerevisiae*
C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 19-Apr-2002
C:Accession: S61990; S66871; S72139
R:Stierky, F.; Uhlen, M.
Submitted to the EMBL Data Library, December 1995
A:Reference number: S61991
A:Accession: S61990
A:Molecule type: DNA
A:Residues: 1-313 <STE>
A:Cross-references: EMBL:U043491; NID:g1150992; PID:g1151002

R:Pettersson, B.; Stierky, F.; Uhlen, M.
Submitted to the Protein Sequence Database, July 1996
A:Reference number: S66882
A:Accession: S66871
A:Molecule type: DNA
A:Residues: 1-313 <PET>
A:Cross-references: EMBL:Z74914; NID:g1420097; PID:e251942; PID:g1420098; MIPS:YOR006
A:Experimental source: strain S288C
R:Stierky, F.; Holmberg, A.; Pettersson, B.; Uhlen, M.
Yeast 12, 1091-1095, 1996
A:Title: The sequence of a 30 kb fragment on the left arm of chromosome XV from *Sacch*
A:Reference number: S72130; MUID:97051599; PMID:8896276
A:Accession: S72139
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-313 <STO>
A:Cross-references: EMBL:U043491; NID:g1150992; PIDN:AAC49486.1; PID:g1151002
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995
C:Genetics:
A:Cross-references: SGD:S0005532
A:Map position: 15R
A:Note: YOR006c

Query Match 65.5%; Score 38; DB 2; Length 313;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 EWLKLLKK 10
:|||||:
DB 207 EWLKLEK 215

RESULT 30
F86462
AGI1-like protein, 41133-42535 [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F86462
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon Chinn, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzia Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tello ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A6141; MUID:21016719; PMID:11130712
A:Accession: F86462
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <STO>
A:Cross-references: GB:AE005172; NID:g6957512; PIDN:AAF32434.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 65.5%; Score 38; DB 2; Length 334;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 EWLKLLKK 10
|||||:
DB 223 EWLKOKK 230

RESULT 31
T12745
hypothetical protein 29 - *Methanobacterium phage psiM2*
C:Species: *Methanobacterium phage psiM2*
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 24-Sep-1999
C:Accession: T12745
R:Pfister, P.; Wasserfallen, A.; Stettler, R.; Leisinger, T.

submitted to the EMBL Data Library, May 1998
A:Description: Archaeophaga Psim2 complete genomic DNA.
A:Reference number: 217578
A:Accession: T12745
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-402 <PFI>
A:Cross-references: EMBL:AF065411; NID:g3249585; PID:g3249614; PIDN:AMC27068.1
A:Experimental source: host Methanobacterium thermoautotrophicum strain Marburg

Query Match 65.5% Score 38; DB 2; Length 402;
Best Local Similarity 54.5% Pred. No. 1.3e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKKLKKM 11
|:|:|:|:|:|:
Db 377 LREWVKTKTKI 387

RESULT 32
F88642
Protein C54E4.3 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: F88642
R:Anonymous, The C. elegans Sequencing Consortium.
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
Science 282, 2012-2018, 1998
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see web sites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_elegans/; published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: F88642
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-583 <SNO>
A:Cross-references: GB:chr.IV; PINN:AAB92031.1; PID:g2702405; GSPDB:GN00022; CESP:C54E4.
A:Note: Similar to guanylate cyclase
C:Genetics:
A:Gene: C54E4.3
A:Map position: 4

Query Match 65.5% Score 38; DB 2; Length 583;
Best Local Similarity 63.6% Pred. No. 1.9e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKKLKKM 11
|:|:|:|:|:|:
Db 376 LRETKMKMKM 386

RESULT 33
J00474
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) STR2 - yeast (Saccharomyces diastaticus)
N:Alternate names: glucanase II
C:Species: Saccharomyces diastaticus
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: J00474; S35895
R:Lambrechts, M.G.; Pretorius, I.S.; Sollitt, P.; Marmur, J.
Gene 100, 95-103, 1991
A:Title: Primary structure and regulation of a glucanase-encoding gene (STR2) in Sac
A:Reference number: J00474; MUID:91276266; PMID:2055484
A:Accession: J00474
A:Molecule type: DNA
A:Residues: 1-767 <LJM>

A:Cross-references: GB:M60650; NID:g172733; PINN:AAA35107.1; PID:g172734
A:Note: the authors translated the codon CCG for residue 337 as Ala, CCG for residue 364
R:Kum, K.; Bayszar, G.; Lee, S.Y.; Knudsen, F.; Mattoon, J.R.
Submitted to the EMBL Data Library, May 1992
A:Description: Cloning of a new allelic variant of a Saccharomyces diastaticus glucanase
A:Reference number: S35895
A:Accession: S35895
A:Molecule type: DNA
A:Residues: 1-163, 'A', 164-622, 'D', 624-767 <KIM>

A:Cross-references: EMBL:M90490; NID:g172735; PIDN:AAA20560.1; PID:g172736
C:Genetics:
A:Gene: STR2; DEX1
C:Superfamily: yeast glucan 1,4-alpha-glucosidase STR1; glucan 1,4-alpha-glucosidase
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:334-767/Domain: glucan 1,4-alpha-glucosidase homology <GAG>

Query Match 65.5% Score 38; DB 1; Length 767;
Best Local Similarity 54.5% Pred. No. 2.5e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKKLKKM 11
|:|:|:|:|:|:
Db 338 LREWLQROKKV 348

RESULT 34
ALBYG
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - yeast (Saccharomyces cerevisiae)
N:Alternate names: 1,4-alpha-D-glucan glucosylhydrolase; glucanase
C:Species: Saccharomyces cerevisiae
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 05-Sep-1997
C:Accession: A21896; A23470
R:Yamashita, I.; Suzuki, K.; Fukui, S.
J. Bacteriol. 161, 567-573, 1985
A:Title: Nucleotide sequence of the extracellular glucanase gene STR1 in the yeast
A:Reference number: A21896; MUID:85104778; PMID:3918017
A:Note: S. diastaticus
A:Accession: A21896
A:Molecule type: DNA
A:Residues: 1-778 <YAI>
A:Experimental source: strain 5106-9A; ATCC 60709
R:Yamashita, I.; Suzuki, K.; Sakuzo, F.
Agric. Biol. Chem. 50, 475-482, 1986
A:Title: Polyoxytic processing of glucoamylase in the yeast Saccharomyces cerevisiae
A:Reference number: A23470
A:Contents: signal sequence cleavage site
A:Accession: A23470
A:Molecule type: protein
A:Residues: 1-65 <YAZ>
C:Genetics:
A:Gene: STR1
C:Superfamily: yeast glucan 1,4-alpha-glucosidase STR1; glucan 1,4-alpha-glucosidase
C:Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; polysaccharat
F:1-32/Domain: signal sequence #status experimental <SIG>
F:33-778/Product: glucan 1,4-alpha-glucosidase #status predicted <MP7>
F:345-778/Domain: glucan 1,4-alpha-glucosidase homology <GAG>
F:46,319,333,425,434,445,524,557,656,661,731,752/Binding site: carbohydrate (Asn) (Co
F:635/Active site: Asp #status predicted

Query Match 65.5% Score 38; DB 1; Length 778;
Best Local Similarity 54.5% Pred. No. 2.5e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKKLKKM 11
|:|:|:|:|:|:
Db 349 LREWLQROKKV 359

RESULT 35
H69980
single-strand DNA-specific exonuclease homolog yve - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C:Accession: H69980
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Bruliet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Erlington, J.; Fabret, C.; Ferraril,
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galliz, A.; Gal
teich, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau

Y, M.; Ogawa, K.; Ogiwara, A.; Oudegaa, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, A.; Scroffone, F.; Sekiguchi, J.; Sekowska, A.; Serot, A.; Winters, P.; Wipac, A.; Tanaka, T.; Terstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A. A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. A:Reference number: A69580; MUID:98044033; PMID:9384377 A:Accession: H69980 A:Status: preliminary; nucleic acid sequence not shown; translation not shown A:Molecule type: DNA A:Residues: 1-786 <KUN> A:Cross-references: GB:259118; GB:AL009126; NID:g2635200; PIDN:CAB14721.1; PID:e1184010; A:Experimental source: strain 168 C:Genetics: A:Gene: yrvE

Query Match 65.5%; Score 38; DB 2; Length 786; Best Local Similarity 70.0%; Pred. No. 2.6e+02; Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKKLK 10
|:|:|:|:|:|
Db 765 LKEWLKKLK 774

RESULT 36
T21085
hypothetical protein F18C12.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T21085
R:Harits, B.
Submitted to the EMBL Data Library, June 1996
A:Reference number: Z19371
A:Accession: T21085
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-4131 <WIL>
A:Cross-references: EMBL:Z75536; PIDN:CAA98830.1; GSPDB:GN00019; CESP:F18C12.1
C:Genetics: A:Experimental source: clone F18C12
A:Gene: CESP:F18C12.1
A:Map position: 1
A:Intons: 54/3; 112/3; 146/1; 294/1; 381/3; 455/3; 596/3; 716/3; 743/2; 782/3; 820/3; 809/3; 2303/3; 2483/1; 2529/1; 2612/3; 2687/1; 2734/3; 2799/3; 2841/3; 2891/2; 2932/3; 29 C:Superfamily: dynein heavy chain, cytosolic

Query Match 65.5%; Score 38; DB 2; Length 4131; Best Local Similarity 66.7%; Pred. No. 1.4e+03; Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOEWLKKLK 9
|:|:|:|:|:|
Db 1124 LQKMKLK 1132

RESULT 37
AG1740
B. subtilis general stress protein 13 containing a ribosomal S1 protein domain homolog 1
C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AG1740
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1740
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-117 <GUA>
A:Cross-references: GB:AL592022; PIDN:CA97695.1; PID:g16414990; GSPDB:GN00178
A:Experimental source: strain C11p11262
C:Genetics: A:Gene: lln2468
C:Superfamily: polyribonucleotide nucleotidyltransferase homolog yabr

Query Match 63.8%; Score 37; DB 2; Length 117; Best Local Similarity 60.0%; Pred. No. 53; Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKKLK 10
|:|:|:|:|:|
Db 108 LEEWIKKADK 117

RESULT 38
A11370
B. subtilis general stress protein 13 containing a ribosomal S1 protein domain homolog 1
C:Species: *Listeria monocytogenes*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: A11370
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: A11370
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-117 <GUA>
A:Cross-references: GB:NC_003210; PIDN:CAD00447.1; PID:g16411857; GSPDB:GN00177
C:Genetics: A:Experimental source: strain EGD-e
A:Gene: lmo2369
C:Superfamily: polyribonucleotide nucleotidyltransferase homolog yabr

Query Match 63.8%; Score 37; DB 2; Length 117; Best Local Similarity 60.0%; Pred. No. 53; Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKKLK 10
|:|:|:|:|:|
Db 108 LEEWIKKADK 117

RESULT 39
B87702
ribosomal protein S16 [imported] - *Caulobacter crescentus*
C:Species: *Caulobacter crescentus*
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: B87702
R:Nierman, W.C.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwynn, M.L.; Haft, D.H.; Koh, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: B87702
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-165 <STO>
A:Cross-references: GB:AE005673; NID:g13425408; PIDN:AAK25614.1; GSPDB:GN00148
C:Genetics: A:Gene: CC3652

Query Match 63.8%; Score 37; DB 2; Length 165; Best Local Similarity 85.7%; Pred. No. 75; Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOEWLKK 7
:|||||
DB 58 IOEWLKK 64

RESULT 40

C64366
hypothetical protein homolog MJ0531 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: C64366
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Rison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kalne, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: C64366
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-170 <BUL>
A:Cross-references: GB:067502; GB:L77117; NID:q2826293; PIDN:AA898522.1; PID:q1591234; T
C:Genetics:
A:Map position: FOR467095-467607
C:Superfamily: Escherichia coli ybdQ protein

Query Match 63.8%; Score 37; DB 2; Length 170;
Best Local Similarity 80.0%; Pred. No. 78;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 QEWLKK 11
:|||||
DB 89 QEWLKK 98

RESULT 41

T05164

hypothetical protein F18E5.180 - Arabidopsis thaliana

N:Alternate names: hypothetical protein F17L22.20
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C:Accession: T05164; T05831
R:Bayan, M.; Peters, S.A.; Van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mey
submitted to the Protein Sequence Database, August 1998
A:Reference number: Z15400
A:Accession: T05164
A:Molecule type: DNA
A:Residues: 1-209 <BEV>
A:Cross-references: EMBL:AL022603
A:Experimental source: cultivar Columbia; BAC clone F18E5
R:Bayan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ba
submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15454
A:Accession: T05831
A:Molecule type: DNA
A:Residues: 1-209 <BEV>
A:Cross-references: EMBL:AL035527
A:Experimental source: cultivar Columbia; BAC clone F17L22
C:Genetics:
A:Map position: 4
A:Note: F18E5.180; F17L22.20

Query Match 63.8%; Score 37; DB 2; Length 209;
Best Local Similarity 54.5%; Pred. No. 96;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKK 11
:|||||
DB 165 MKEWLKK 175

RESULT 42

F72272

conserved hypothetical protein - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: F72272
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: F72272
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-218 <ARN>
A:Cross-references: GB:AE001784; GB:AE000512; NID:q4981842; PIDN:AD36369.1; PID:q498
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1295
C:Superfamily: glyoxalase

Query Match 63.8%; Score 37; DB 2; Length 218;
Best Local Similarity 63.6%; Pred. No. 1e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKK 11
:|||||
DB 162 LREYLSLKL 172

RESULT 43

C70380

hypothetical protein aq_926 - Aquifex aeolicus

C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: C70380
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: C70380
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-252 <ADQ>
A:Cross-references: GB:AE000714; NID:q2983446; PIDN:AA07036.1; PID:q2983455; GB:AE00
A:Experimental source: strain VFS
C:Genetics:
A:Gene: aq_926

Query Match 63.8%; Score 37; DB 2; Length 252;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKK 10
:|||||
DB 208 LKSNKK 217

RESULT 44

AE2237

hypothetical protein alr3452 (imported) - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AE2237
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kurita, T.; Sasamoto, S.; Watanabe, A.; Iriku
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genome Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AE2237

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-415 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA075151.1; PID:g17132585; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:

Query Match 63.8%; Score 37; DB 2; Length 415;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LQEWLKK 7
|:|||||
Db 380 MQEWLKK 386

RESULT 45

XNECDP
adenosylmethionine-8-amino-7-oxononanoate transaminase (EC 2.6.1.62) b10A [similarity]
N:Alternate names: 7,8-diaminononanoate aminotransferase; adenosylmethionine-8-amino-7-C
C:Species: Escherichia coli
C:Date: 30-Jun-1990 #sequence_revision 17-Oct-1997 #text_change 03-Jun-2002
C:Accession: F64813; A32025
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: F64813
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-429 <BLAT>
A:Cross-references: GB:AE000180; GB:U00096; NID:g1786988; PIDN:AAC73861.1; PID:g1786991;
A:Experimental source: strain K-12, substrain MG1655
R:Otsuka, A.J.; Bonocristiani, M.R.; Howard, P.K.; Flamm, J.; Johnson, C.; Yamamoto, R.;
J. Biol. Chem. 263, 19577-19585, 1988
A:Title: The Escherichia coli biotin biosynthetic enzyme sequences predicted from the nu
A:Reference number: A32025; MUID:89066784; PMID:3058702
A:Accession: A32025
A:Molecule type: DNA

A:Residues: 1-98, 'SGRNA', 103-429 <OTS>
A:Cross-references: GB:J04423; NID:g145422; PIDN:AAA3514.1; PID:g457106
C:Genetics:
A:Gene: b10A
A:Map position: 17 min
C:Complex: homodimer
C:Superfamily: ornithine-oxo-acid aminotransferase
C:Keywords: aminotransferase; biotin biosynthesis; homodimer; phosphoprotein; pyridoxal
P;274/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 63.8%; Score 37; DB 1; Length 429;
Best Local Similarity 55.6%; Pred. No. 2e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 3 EWLKLRK 11
|:|||||
Db 226 EWLKLRK 234

RESULT 46

T42366
L-aminodipate-semialdehyde dehydrogenase homolog - fission yeast (Schizosaccharomyces f
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: T42366
R:Yoshida, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
A:Reference number: 217323; MUID:98162722; PMID:9501991
A:Accession: T42366
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

A:Residues: 1-445 <YOS>
A:Cross-references: EMBL:D89124; NID:g1749455; PIDN:BA013786.1; PID:g1749456
A:Experimental source: strain PR745

Query Match 63.8%; Score 37; DB 2; Length 445;
Best Local Similarity 60.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 LQEWLKK 10
|:|||||
Db 378 LEWMLALEK 387

RESULT 47

AI2249
exopolysphatase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AI2249
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AI2249
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-550 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA075251.1; PID:g17132685; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all3552
C:Superfamily: exopolysphatase

Query Match 63.8%; Score 37; DB 2; Length 550;
Best Local Similarity 45.5%; Pred. No. 2.6e+02;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Oy 1 LQEWLKK 11
|:|||||
Db 266 LEWVNRLRM 276

RESULT 48

D82910
aspartyl-tRNA synthetase U0286 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: D82910
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a
A:Reference number: A82870
A:Accession: D82910
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-578 <GLA>
A:Cross-references: GB:AE00125; GB:AF222894; NID:96899253; PIDN:AAF30695.1; GSPDB:GN
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: aspS; U0286
A:Genetic code: SCC3
C:Superfamily: lysine-tRNA ligase

Query Match 63.8%; Score 37; DB 2; Length 578;
Best Local Similarity 54.5%; Pred. No. 2.7e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 LQEWLKK 11
|:|||||
Db 19 LMGVKKVRKM 29

RESULT 49

T24458

hypotheetical protein T04D3.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C:Accession: T24458

R:Kershaw, J.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19893

A:Accession: T24458

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-751 <WIL>

A:Cross-references: EMBL:Z81114; PIDN:CA803288.1; GSPDB:GN00019; CESP:T04D3.4

A:Experimental source: clone T04D3

C:Genetics:

A:Gene: CESP:T04D3.4

A:Map position: 1

A:Introns: 1/3; 26/2; 95/3; 149/1; 213/3; 247/3; 291/3; 332/3; 369/2; 418/1; 462/3; 498/

C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

Query Match

63.8%; Score 37; DB 2; Length 751;

Best Local Similarity 72.7%; Pred. No. 3.5e+02;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKLKKM 11

|:| |11111

DB 372 LEETTKLKKM 382

RESULT 50

AD3232

conjugal transfer protein trbE [imported] - Agrobacterium tumefaciens (strain C58, Dupon

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 17-May-2002

C:Accession: AD3232

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo, I

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavln, T.; Levy, R.; Li, M.; McCell

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AD3232

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-822 <KUR>

A:Cross-references: GB:AE008690; PIDN:AAL46274.1; PID:q17744055; GSPDB:GN00189

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: trbE

A:Genome: plasmid

C:Superfamily: virB4 protein

Query Match

63.8%; Score 37; DB 2; Length 822;

Best Local Similarity 60.0%; Pred. No. 3.9e+02;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKLKK 10

::111111

DB 678 IREWLKVLRK 687

Search completed: July 11, 2003, 10:49:51

Job time : 51 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 10:43:18 ; Search time 23 Seconds
(without alignments)
19,837 Million cell updates/sec

Title: US-09-828-592-10
Perfect score: 58
Sequence: 1 LQEWLKKKKM 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 300 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	72.4	509	1	TSSP_MOUSE
2	40	69.0	464	1	ANT3_HUMAN
3	40	69.0	1323	1	RRPO_SWINE
4	39	67.2	207	1	P1G1_MOUSE
5	39	67.2	209	1	V281_ARATH
6	39	67.2	928	1	DPOL_SALTY
7	39	67.2	1290	1	P1G1_HUMAN
8	39	67.2	1290	1	P1G1_MOUSE
9	39	67.2	1291	1	P1G1_BOVIN
10	38	65.5	98	1	ACTO_CHICK
11	38	65.5	313	1	Y006_YEAST
12	38	65.5	767	1	AMVH_HUMAN
13	38	65.5	768	1	AMVH_SACDI
14	38	65.5	1015	1	YAI7_HUMAN
15	37.5	64.7	231	1	Y339_AOUAE
16	37	63.8	165	1	RS16_CAVUR
17	37	63.8	170	1	Y531_METJA
18	37	63.8	209	1	V282_ARATH
19	37	63.8	429	1	B10A_ECOLI
20	37	63.8	485	1	G1CA_HUMAN
21	37	63.8	662	1	TRB2_RHISN
22	37	63.8	820	1	TRBE_AGRU
23	37	63.8	964	1	MSH2_YEAST
24	37	63.8	1070	1	EX5C_BUCAL
25	37	63.8	1643	1	RRPO_MOUSE
26	36	62.1	161	1	XKDI_BACSU
27	36	62.1	200	1	COAE_MYCPN
28	36	62.1	223	1	RB21_CANPA
29	36	62.1	324	1	RB21_HUMAN
30	36	62.1	374	1	2NPD_WILMR
31	36	62.1	423	1	SYT_ECOLI
32	36	62.1	438	1	ODP2_NEUCR
33	36	62.1	492	1	MS_STRPS

34	36	62.1	551	1	VG28_BPMU
35	36	62.1	632	1	PF22_CAEEL
36	36	62.1	646	1	PIXB_HUMAN
37	36	62.1	691	1	Y104_YEAST
38	36	62.1	1056	1	AMP1_PLAFO
39	36	62.1	2414	1	P300_HUMAN
40	36	62.1	2441	1	CBP_MOUSE
41	36	62.1	2442	1	CBP_HUMAN
42	35	60.3	63	1	R24E_AERPE
43	35	60.3	140	1	CEN4_HUMAN
44	35	60.3	231	1	YE22_YEAST
45	35	60.3	264	1	KGUA_MYGE
46	35	60.3	264	1	NNMT_MOUSE
47	35	60.3	264	1	TEMT_MOUSE
48	35	60.3	353	1	AROC_BUCAL
49	35	60.3	406	1	IPSP_HUMAN
50	35	60.3	497	1	YCA6_DDOST
51	35	60.3	564	1	YEDQ_ECOS7
52	35	60.3	564	1	YEDQ_ECOLI
53	35	60.3	776	1	ARH6_HUMAN
54	35	60.3	793	1	PHK2_ANASP
55	35	60.3	796	1	PHK_CLOAB
56	35	60.3	1036	1	YAN2_SCHRO
57	35	60.3	1067	1	CARB_CLOPE
58	35	60.3	1574	1	MY52_YEAST
59	35	60.3	1718	1	RRPO_SHVX
60	34	58.6	66	1	CART_PIG
61	34	58.6	84	1	NR9M_HUMAN
62	34	58.6	116	1	CART_HUMAN
63	34	58.6	129	1	CART_MOUSE
64	34	58.6	129	1	CART_RAT
65	34	58.6	191	1	CBX5_HUMAN
66	34	58.6	191	1	CBX5_MOUSE
67	34	58.6	194	1	RB31_HUMAN
68	34	58.6	196	1	GCN2_RANPL
69	34	58.6	197	1	GCN2_CHICK
70	34	58.6	199	1	GCN2_HUMAN
71	34	58.6	203	1	GCN2_BOVIN
72	34	58.6	228	1	CLCB_BOVIN
73	34	58.6	229	1	CLCB_HUMAN
74	34	58.6	229	1	CLCB_RAT
75	34	58.6	229	1	YMS1_ARCFU
76	34	58.6	316	1	MTRH_METMA
77	34	58.6	327	1	GRHR_MOUSE
78	34	58.6	385	1	YG1M_YEAST
79	34	58.6	393	1	P2CA_YEAST
80	34	58.6	416	1	RPSD_MICAE
81	34	58.6	446	1	YDB5_PYROO
82	34	58.6	459	1	GLMU_BUCAL
83	34	58.6	465	1	ANT3_MOUSE
84	34	58.6	471	1	RNT1_YEAST
85	34	58.6	486	1	DNAB_RICPR
86	34	58.6	503	1	AGN4_HUMAN
87	34	58.6	517	1	GUNA_CLOLO
88	34	58.6	549	1	AMVG_YEAST
89	34	58.6	550	1	SYD_MYGE
90	34	58.6	564	1	ROK1_YEAST
91	34	58.6	570	1	CRT1_SOYRN
92	34	58.6	645	1	GUP1_YEAST
93	34	58.6	934	1	MSH2_HUMAN
94	34	58.6	1024	1	CARC_HUMAN
95	34	58.6	1089	1	CARB_CAMJE
96	34	58.6	1161	1	DP3A_AOUAE
97	34	58.6	1385	1	RRPO_PLANY
98	34	58.6	1513	1	STU1_YEAST
99	34	58.6	1547	1	RRPO_PMV
100	34	58.6	1663	1	HAPP_HUMAN
101	34	58.6	1919	1	HAPP_RAT
102	34	58.6	4563	1	APB_HUMAN
103	33.5	57.8	153	1	Y1BG_ECOLI
104	33	56.9	148	1	ETP_ECOLI
105	33	56.9	164	1	RB13_RAT
106	33	56.9	203	1	RB13_HUMAN

09C1W6 bacterioph
009312 caenorhabd
014155 homo sapien
P40460 saccharomyc
096935 plasmodium
009472 homo sapien
P45481 mus musculu
092793 homo sapien
09Y947 aeropyrum p
P49450 homo sapien
P40007 saccharomyc
P47353 mycoplasma
055239 mus musculu
P40936 mus musculu
092469 buchnera ap
P05154 homo sapien
P45451 odontella s
08X892 escherichia
P76330 escherichia
015052 homo sapien
08Y126 anabaena sp
0971e3 clostridium
010068 schizosacch
08X8b3 clostridium
P19524 saccharomyc
004575 shalilot vlr
095167 homo sapien
095167 homo sapien
P15588 mus musculu
P49192 rattus norv
P45973 homo sapien
061686 mus musculu
013635 homo sapien
073762 rana pipien
P79881 gallus galli
09umx6 homo sapien
P51177 bos taurus
P04975 bos taurus
P09497 homo sapien
P08082 rattus norv
028033 archaeglob
001650 methanosarc
P01776 mus musculu
P53230 saccharomyc
P36089 saccharomyc
P52222 microcystis
050093 pyrococcus
P57139 buchnera ap
002555 saccharomyc
094d08 rickettsia
092644 homo sapien
P34937 clostridium
P08019 saccharomyc
P47282 mycoplasma
P48818 glycoline max
P46943 saccharomyc
P43246 homo sapien
09np44 homo sapien
096117 campylobact
061125 aquiflex aeo
007518 plantago as
P38198 saccharomyc
P20951 papaya mosa
060229 homo sapien
P79924 rattus norv
P04114 homo sapien
P32106 escherichia
P75880 escherichia
P35286 rattus norv
P51153 homo sapien

107	33	56.9	213	1	RABR_DICDI	P34142 dictyosteli	180	32	55.2	226	1	YC65_MYCTU	Q11054 mycobacteri
108	33	56.9	235	1	GTL_ONCVO	P46634 onchocerca	181	32	55.2	231	1	MAG1_PETHY	Q07472 lactonia hyb
109	33	56.9	244	1	YENR_YEREN	P54295 yersinia en	182	32	55.2	232	1	ABGX_IATJC	Q48724 lactococcus
110	33	56.9	258	1	Y144_METUA	O57608 methanococ	183	32	55.2	238	1	YEBN_ECOLI	P76484 escherichia
111	33	56.9	266	1	MA33_YEAST	P40313 saccharomyc	184	32	55.2	242	1	TRP2_CYACA	P34793 cyanidium c
112	33	56.9	269	1	SPEE_ARCFU	O27950 archaeoglob	185	32	55.2	243	1	TRPA_CYAME	O22018 cyanidiosch
113	33	56.9	297	1	RL5_HELAN	O65353 heilanthus	186	32	55.2	253	1	HIS6_THETN	O81885 thermoaer
114	33	56.9	300	1	MNTA_BACHD	O9K673 bacillus ha	187	32	55.2	255	1	RNH2_BACSU	O31744 bacillus su
115	33	56.9	319	1	ANIL_EMENT	P03880 emeticella	188	32	55.2	263	1	INMT_HUMAN	O95050 homo sapien
116	33	56.9	322	1	Y711_METUA	O58121 methanococ	189	32	55.2	264	1	NNMT_HUMAN	P40261 homo sapien
117	33	56.9	327	1	GRHR_RAT	P30969 rattus norv	190	32	55.2	270	1	TRPA_BUCDN	O68429 buchnera ap
118	33	56.9	343	1	BIOB_BUCAT	P57378 buchnera ap	191	32	55.2	282	1	DAAA_BACSP	P19938 bacillus sp
119	33	56.9	368	1	Y812_AQUAE	O66994 aquifex ae	192	32	55.2	283	1	DAAA_BACSH	P54693 bacillus sp
120	33	56.9	369	1	TP6A_METUA	O57815 methanococ	193	32	55.2	283	1	PDXK_ECOLI	P40191 escherichia
121	33	56.9	373	1	RFL_AERPE	O9VAF1 aeropyrum p	194	32	55.2	295	1	XERC_HAEIN	P44818 haemophilus
122	33	56.9	405	1	S3AE_BACSU	P16372 mus musculu	195	32	55.2	311	1	PP1L_YEAST	P20604 saccharomyc
123	33	56.9	411	1	RAPS_MOUSE	O58383 methanococ	196	32	55.2	314	1	TPIC_FFRAN	O9M458 fragaria an
124	33	56.9	411	1	Y973_METUA	O52250 halomonas e	197	32	55.2	329	1	YOG1_CAEEL	O09272 caenorhabd
125	33	56.9	421	1	ECB2_HAIEL	O9UKS6 homo sapien	198	32	55.2	331	1	ISPH_LISMO	P58676 listeria mo
126	33	56.9	424	1	PAC3_HUMAN	O991b8 mus musculu	199	32	55.2	331	1	MAN1_MOUSE	O9WU40 mus musculu
127	33	56.9	424	1	PAC3_MOUSE	P00506 sus scrofa	200	32	55.2	350	1	Y555_METJA	O57975 methanococ
128	33	56.9	430	1	AATM_PIG	P41361 bos taurus	201	32	55.2	351	1	Y239_METJA	O60294 methanococ
129	33	56.9	433	1	ANT3_BOVIN	Q13685 homo sapien	202	32	55.2	353	1	PLE2_HUMAN	O9NYT0 homo sapien
130	33	56.9	452	1	AAMP_HUMAN	P32262 ovis aries	203	32	55.2	353	1	PLE2_MOUSE	O9WV52 mus musculu
131	33	56.9	465	1	ANT3_SHEEP	O66766 aquifex ae	204	32	55.2	355	1	ISPG_THETN	O84303 thermoaer
132	33	56.9	478	1	GATB_AQUAE	P22250 bacillus su	205	32	55.2	358	1	AROG_BACSU	P39912 b aroa(9) p
133	33	56.9	483	1	SYE_BACSU	P23552 caldocellum	206	32	55.2	361	1	RL3_PYRAB	O9Y1T5 pyrococcus
134	33	56.9	488	1	XYNB_CALSA	P79739 brachydanio	207	32	55.2	362	1	RL3_PYRHO	O54048 pyrococcus
135	33	56.9	492	1	CP26_BRARE	O08607 mycobacteri	208	32	55.2	362	1	Y024_MYCPN	P73088 mycoplasma
136	33	56.9	493	1	MOO_MYCTU	P42251 bacillus su	209	32	55.2	385	1	TPST_DROME	O9YAB7 drosophila
137	33	56.9	556	1	PPBD_BACSU	P75068 mycoplasma	210	32	55.2	395	1	KIME_MOUSE	O9Y008 mus musculu
138	33	56.9	557	1	SYD_MYCPN	P36158 saccharomyc	211	32	55.2	399	1	CYH3_MOUSE	O08967 mus musculu
139	33	56.9	585	1	YK58_YEAST	P26996 thermus the	212	32	55.2	400	1	CYH3_HUMAN	O43739 homo sapien
140	33	56.9	676	1	DNLJ_THETH	O27597 drosophila	213	32	55.2	400	1	CYH3_RAT	P97696 rattus norv
141	33	56.9	679	1	NCPR_DROME	P40310 saccharomyc	214	32	55.2	406	1	O10A_DROME	O9YVZ1 drosophila
142	33	56.9	691	1	TOK1_YEAST	P22956 red clover	215	32	55.2	416	1	Y997_ARCFU	O23265 archaeoglob
143	33	56.9	767	1	RPOJ_RCNMV	P34603 caenorhabd	216	32	55.2	422	1	RPC4_YEAST	P25441 saccharomyc
144	33	56.9	784	1	Y063_CAEEL	P33811 homo sapien	217	32	55.2	435	1	RUBB_CHLRE	O42693 chlamydomon
145	33	56.9	841	1	TRK_HUMAN	P47326 mycoplasma	218	32	55.2	465	1	NP11_MOUSE	P461983 mus musculu
146	33	56.9	848	1	OPPE_MYCPE	P70755 aquifex pyr	219	32	55.2	465	1	Y065_MYCPN	P75612 mycoplasma
147	33	56.9	855	1	MURS_AOUCY	P35761 mus musculu	220	32	55.2	466	1	Y065_MYCCE	P47311 mycoplasma
148	33	56.9	856	1	TRK_MOUSE	P42173 caenorhabd	221	32	55.2	471	1	SG3_MOUSE	P47867 mus musculu
149	33	56.9	922	1	YK16_CAEEL	O57588 methanococ	222	32	55.2	471	1	SG3_RAT	P47868 rattus norv
150	33	56.9	1075	1	Y124_METUA	P08548 nyclicebus	223	32	55.2	473	1	PHR_SALTY	P25078 salmoneilla
151	33	56.9	1249	1	METH_CAEEL	P11978 saccharomyc	224	32	55.2	485	1	TR2_SYNY3	P74130 synecocyst
152	33	56.9	1260	1	LINI_MYCCO	P78595 candida alb	225	32	55.2	487	1	IA11_ORYSA	O07215 oryza sativ
153	33	56.9	1358	1	SIR4_YEAST	O15078 homo sapien	226	32	55.2	492	1	6REG_SCHPO	P78812 schizosach
154	33	56.9	1499	1	CDR2_CANAL	P34345 caenorhabd	227	32	55.2	498	1	YDHF_SCHPO	P78334 homo sapien
155	33	56.9	1539	1	Y373_HUMAN	O62261 mus musculu	228	32	55.2	506	1	GAE_HUMAN	P46571 buchnera ap
156	33	56.9	2056	1	CBP1_CAEEL	P02549 homo sapien	229	32	55.2	518	1	LEU1_BUCRP	P34492 caenorhabd
157	33	56.9	2363	1	SPCO_MOUSE	O01082 homo sapien	230	32	55.2	526	1	WR11_ARATH	O93470 arabidopsis
158	33	56.9	2364	1	SPCA_HUMAN	O05002 human corion	231	32	55.2	538	1	REC8_SCHPO	P36626 schizosach
159	33	56.9	2418	1	RRPA_CVH22	P30597 oryctolagus	232	32	55.2	557	1	PYRH_YEAST	P36627 schizosach
160	33	56.9	4085	1	RIR2_RABIT	P47763 yersinia en	233	32	55.2	617	1	IF2P_AERPE	O9Y933 aeropyrum p
161	33	56.9	4969	1	MURS_YEREN	P45012 haemophilus	234	32	55.2	630	1	TFC_MOUSE	P24604 mus musculu
162	32	55.2	38	1	CCEC_DROME	P13620 bos taurus	235	32	55.2	631	1	TFC_HUMAN	P42680 homo sapien
163	32	55.2	63	1	FABZ_HAEIN	O58346 methanococ	236	32	55.2	631	1	Y053_HUMAN	P43331 homo sapien
164	32	55.2	148	1	ATPO_BOVIN	P36540 escherichia	237	32	55.2	646	1	YEL4_SCHPO	O13865 schizosach
165	32	55.2	160	1	Y936_METUA	P36540 escherichia	238	32	55.2	655	1	AMTA_PYRAB	O92298 pyrococcus
166	32	55.2	166	1	HSCB_ECOLI	P38879 enterobacte	239	32	55.2	674	1	AGUA_THEMEA	P96105 thermotoga
167	32	55.2	171	1	FLAV_ENTAG	O33787 beta trigen	240	32	55.2	697	1	DN1J_THESK	O94FY8 thermus sp.
168	32	55.2	177	1	NUGM_BETR	O34011 beta webbia	241	32	55.2	705	1	PKR_BACHD	O99421 dictyosteli
169	32	55.2	192	1	NUGM_BETVU	O9Z4d8 buchnera ap	242	32	55.2	714	1	CLPB_MYCCE	P47597 mycoplasma
170	32	55.2	192	1	NUGM_BETWE	P34362 caenorhabd	243	32	55.2	769	1	TOP1_YEAST	P04786 saccharomyc
171	32	55.2	193	1	DCD_BUCAP	O9H0C7 homo sapien	244	32	55.2	777	1	PPB2_VIBCH	O9WUC0 v penicilli
172	32	55.2	196	1	YLH8_CAEEL	P35292 mus musculu	245	32	55.2	779	1	PKR2_RHIME	O922a4 rhizobium m
173	32	55.2	212	1	RL17_HUMAN	P39811 escherichia	246	32	55.2	807	1	Y066_NPVAC	O98467 autographa
174	32	55.2	214	1	GLR2_ECOLI	O82d91 yersinia pe	247	32	55.2	808	1	RSG5_HUMAN	O43374 homo sapien
175	32	55.2	215	1	APP6_YERPE	P41291 balaeopter	248	32	55.2	826	1	TRNL_YEAST	P09880 saccharomyc
176	32	55.2	220	1	APP6_BALPH	P24945 balaeopter	249	32	55.2	874	1	SYLM_SCHPO	O09828 schizosach
177	32	55.2	226	1	ATP6_BALPH		250	32	55.2				
178	32	55.2	226	1	ATP6_BALPH		251	32	55.2				
179	32	55.2	226	1	ATP6_BALPH		252	32	55.2				

ID	TSSP_MOUSE	STANDARD:	PRT:	509 AA.
253				
254				
255				
256				
257				
258				
259				
260				
261				
262				
263				
264				
265				
266				
267				
268				
269				
270				
271				
272				
273				
274				
275				
276				
277				
278				
279				
280				
281				
282				
283				
284				
285				
286				
287				
288				
289				
290				
291				
292				
293				
294				
295				
296				
297				
298				
299				
300				

ALIGNMENTS

RESULT 1

TSSP_MOUSE STANDARD: PRT: 509 AA.

AC 090XES:

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Thymus-specific serine protease precursor (EC 3.4.-.-).

CN PRS16 OR TSSP.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Embryonic thymus;

RX MEDLINE=20093985; PubMed=10630289;

RA Carlier A., Nguyen C., Victorero G., Granjeaud S., Rocha D.,

RA Bernard K., Mlizek A., Ferrier P., Malissen M., Naquet P.,

RA Malissen B., Jordan B.R.;

RT "Differential gene expression in Cd3epsilon- and Rag1-deficient thymuses: definition of a set of genes potentially involved in T lymphocyte maturation."

RT Immunogenetics 50:255-270(1999).

RL -1- FUNCTION: PROTEASE THAT MAY PLAY A ROLE IN T-CELL DEVELOPMENT.

CC -1- SUBCELLULAR LOCATION: VESICULAR, EITHER LYSOSOMAL OR ENDOPLASMIC.

CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN CORTICAL THYMIC EPITHELIAL CELLS, WITH HIGHEST EXPRESSION AROUND VESSELS AND THE THYMIC CAPSULE.

CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN DEVELOPING THYMUS AT 14 TO 18 DPC, WITH MAXIMAL EXPRESSION AT 16 DPC.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S28.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL: AJ131775; CAB6137.1; -

CC MEROPS: S28 003; -

DR MGD: MGI:1859181; Prss16.

DR InterPro: IPR000379; Ser_estrs_site.

KM Hydrolyase; Serine protease; Signal.

FT SIGNAL 1 22

FT CHAIN 23 509 THYMUS-SPECIFIC SERINE PROTEASE.

FT ACT_SITE 184 184 CHARGE RELAY SYSTEM (POTENTIAL).

FT ACT_SITE 446 446 CHARGE RELAY SYSTEM (POTENTIAL).

FT ACT_SITE 471 471 CHARGE RELAY SYSTEM (POTENTIAL).

FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).

SO SEQUENCE 509 AA; 54523 MW; CD8425F3F71EE43C CRC64;

Query Match 72.4%; Score 42; DB 1; Length 509;

Best Local Similarity 70.0%; Pred. No. 13;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LOEWLKKKK 10

Db 497 LQDWLKKKK 506

RESULT 2

ANT3_HUMAN STANDARD: PRT: 464 AA.

AC P01008:

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Antithrombin-III precursor (A1111) (PRO0309).

CN SERPINC1 OR AT3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=83143280; PubMed=6298709;

RX Bock S.C., Wion K.L., Vehar G.A., Lawn R.M.;

RT "Cloning and expression of the cDNA for human antithrombin III."

RT Nucleic Acids Res. 10:8113-8125(1982).

RL [2]

RP SEQUENCE FROM N.A.

RC MEDLINE=83169777; PubMed=6572945;

RX Chandra T., Stackhouse R., Kidd V.J., Woo S.L.C.;

RT "Isolation and sequence characterization of a cDNA clone of human antithrombin III."

RL Proc. Natl. Acad. Sci. U.S.A. 80:1845-1848(1983).

RN [3]

RP SEQUENCE FROM N.A.

RA MEDLINE=93237227; PubMed=8476848;
RA Olds R.J., Lane D.A., Chowdhury V., de Stefano V., Leone G.,
RA Thein S.L.;
RT "Complete nucleotide sequence of the antithrombin gene: evidence for
RT homologous recombination causing thrombophilia.";
RL Biochemistry 32:4216-4224(1993).
RN [14]
RC SEQUENCE FROM N.A.
RC Tissue-Fetal Liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,
RA Xu W., Gao F., Liu M., He F.;
RT "Functional prediction of the coding sequences of 75 new genes deduced
RT by analysis of cDNA clones from human fetal liver.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [15]
RP SEQUENCE FROM N.A., AND VARIANTS GLU-30 AND ALA-147.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [16]
RP SEQUENCE OF 42-464 FROM N.A.
RX MEDLINE=83236456; PubMed=6305982;
RA Prochownik E.V., Markham A.F., Orkin S.H.;
RT "Isolation of a cDNA clone for human antithrombin III.";
RL J. Biol. Chem. 258:8389-8394(1983).
RN [17]
RP SEQUENCE OF 33-464, CARBOHYDRATE-LINKAGE SITES, AND DISULFIDE BONDS.
RA Petersen T.E., Dudek-Wojciechowska G., Sottup-Jensen L.,
RA Magnusson S.;
RT "Primary structure of antithrombin-III (heparin cofactor). Partial
RT homology between alpha-1-antitrypsin and antithrombin-III.";
RL (in) Collen D., Wiman B., Verstraete M. (eds.);
RL The physiological inhibitors of blood coagulation and fibrinolysis,
RL pp.43-54, Elsevier, Amsterdam (1979).
RN [18]
RP ACTIVE SITE.
RX MEDLINE=81212814; PubMed=7238875;
RA Björck I., Danielsson A., Fenton J.W. II, Joernvall H.;
RT "The site in human antithrombin for functional proteolytic cleavage
RT by human thrombin.";
RL FEBS Lett. 126:257-260(1981).
RN [19]
RP HEPARIN-BINDING SITE.
RX MEDLINE=8411578; PubMed=6693405;
RA Blackburn M.N., Smith R.L., Carson J., Sibley C.C.;
RT "The heparin-binding site of antithrombin III. Identification of a
RT critical tryptophan in the amino acid sequence.";
RL J. Biol. Chem. 259:939-941(1984).
RN [10]
RP MUTAGENESIS OF ALA-414.
RX MEDLINE=91192143; PubMed=2013320;
RA Austin R.C., Rachubinski R.A., Blachyjan M.A.;
RT "Site-directed mutagenesis of alanine-382 of human antithrombin III.";
RL FEBS Lett. 280:254-258(1991).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX MEDLINE=94373498; PubMed=8087553;
RA Carrell R.W., Stein P.E., Fernl G., Wardell M.R.;
RT "Biological implications of a 3 A structure of dimeric antithrombin.";
RL Structure 2:257-270(1994).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS).
RX MEDLINE=95384753; PubMed=7656006;
RA Schreuder H.A., de Boer B., Dijkema R., Mulders J., Theunissen H.J.M.,
RA Grootenhuys P.D.J., Hol W.G.J.;
RT "The intact and cleaved human antithrombin III complex as a model for
RT serpin-proteinase interactions.";
RL Nat. Struct. Biol. 1:48-54(1994).
RN [13]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE=97220238; PubMed=9067613;
RA Skinner R., Abraham J.-P., Whistock J.C., Lesk A.M., Carrel R.W.,
RA Wardell M.R.;
RT "The 2.6 A structure of antithrombin indicates a conformational
RT change at the heparin binding site.";
RL J. Mol. Biol. 265:601-609(1997).
RN [14]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE=98437366; PubMed=9761669;
RA Skinner R., Chang W.S.W., Jin L., Pel X.Y., Huntington J.A.,
RA Abraham J.-P., Carrell R.W., Lomas D.A.;
RT "Implications for function and therapy of a 2.9 A structure of
RT binary complexed antithrombin.";
RL J. Mol. Biol. 283:9-14(1998).
RN [15]
RP REVIEW.
RX MEDLINE=91129302; PubMed=2126464;
RA Mourey L., Samama J.-P., Delarue M., Choay J., Lormeau J.C.,
RA Petitou M., Moras D.;
RT "Antithrombin III: structural and functional aspects.";
RL Biochimie 72:599-608(1990).
RN [16]
RP REVIEW ON VARIANTS.
RX MEDLINE=94054329; PubMed=8236149;
RA Lane D.A., Olds R.J., Boiscclair M., Chowdhury V., Thein S.L.,
RA Cooper D.N., Blajchman M., Perry D., Emmerich J., Alach M.;
RT "Antithrombin III mutation database: first update. For the Thrombin
RT and its inhibitors Subcommittee of the Scientific and Standardization
RT Committee of the International Society on Thrombosis and
RT Haemostasis";
RL Thromb. Haemost. 70:361-369(1993).
RN [17]
RP REVIEW ON VARIANTS.
RX MEDLINE=95269065; PubMed=7749926;
RA Stein P.E., Carrell R.W.;
RT "What do dysfunctional serpins tell us about molecular mobility and
RT disease?";
RL Nat. Struct. Biol. 2:96-113(1995).
RN [18]
RP REVIEW ON VARIANTS.
RX MEDLINE=96263733; PubMed=8664906;
RA Perry D.J., Carrell R.W.;
RT "Molecular genetics of human antithrombin deficiency.";
RL Hum. Mutat. 7:7-22(1996).
RN [19]
RP VARIANT UTAH.
RX MEDLINE=89050967; PubMed=3191114;
RA Bock S.C., Marrian J.A., Radziejewska E.;
RT "Antithrombin III Utah: proline-407 to leucine mutation in a highly
RT conserved region near the inhibitor reactive site.";
RL Biochemistry 27:6171-6178(1988).
RN [20]
RP VARIANT TOYAMA.
RX MEDLINE=84119472; PubMed=6582486;
RA Koide T., Odani S., Takahashi K., Ono T., Sakuragawa N.;
RT "Antithrombin III Toyama: replacement of arginine-47 by cysteine in
RT hereditary abnormal antithrombin III that lacks heparin-binding
RT ability.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:289-293(1984).
RN [21]
RP VARIANT BASEL.
RX MEDLINE=86111754; PubMed=3080419;
RA Chang J.Y., Tran T.H.;
RT "Antithrombin III Basel. Identification of a Pro-Leu substitution in
RT a hereditary abnormal antithrombin with impaired heparin cofactor
RT activity.";
RL J. Biol. Chem. 261:1174-1176(1986).
RN [22]
RP VARIANT DENVER.
RX MEDLINE=87109210; PubMed=3805013;
RA Stephens A.W., Thalley B.S., Hirs C.H.W.;
RT "Antithrombin-III Denver, a reactive site variant.";
RL J. Biol. Chem. 262:1044-1048(1987).
RN [23]
RP VARIANT HAMILTON.
RX MEDLINE=89027076; PubMed=3179438;

RA Devyak-Kleuc R., Chui D.H.K., Prochowik E.V., Carter C.J.,
RA Ofosu F.A., Blachman M.A.:
RT "Antithrombin-IIr-Hamilton: a gene with a point mutation (guanine to
RT adenine) in codon 382 causing impaired serine protease reactivity."
RN Blood 72:1518-1523(1988).
[24]
RP VARIANTS GLASCOW AND NORTHWICK-PARK.
RX MEDLINE=88186869; PubMed=3162733;
RA Erdjument H., Lamed D.A., Panico M., di Marzo V., Morris H.R.:
RT "Single amino acid substitutions in the reactive site of antithrombin
RT leading to thrombosis. Congenital substitution of arginine 393 to
RT cysteine in antithrombin Northwick Park and to histidine in
RT antithrombin Glasgow."
RL J. Biol. Chem. 263:5589-5593(1988).
[25]
RN VARIANT CHICAGO.
RP MEDLINE=89388698; PubMed=2781509;
RA Erdjument H., Lane D.A., Panico M., di Marzo V., Morris H.R.,
RA Bauer K., Rosenberg R.D.:
RT "Antithrombin Chicago, amino acid substitution of arginine 393 to
RT histidine."
RL Thromb. Res. 54:613-619(1989).
[26]
RN VARIANT ROUEN-4.
RP MEDLINE=90306344; PubMed=2365065;
RA Borg J.Y., Brennan S.O., Carrell R.W., George P., Perry D.J., Shaw J.:
RT "Antithrombin Rouen-IV 24 Arg-->Cys. The amino-terminal contribution
RT to heparin binding.";

Query Match 69.0%; Score 40; DB 1; Length 464;
Best Local Similarity 63.6%; Pred. No. 25;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 LOEWLKKKKM 11
Db 336 LOEWLDELEEM 346

RESULT 3
ID RPO_SMYEA STANDARD: PRT: 1323 AA.
AC P28897;

DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA replication protein (150 kDa protein) (ORF 1) [Contains: RNA-
DE directed RNA polymerase (EC 2.7.7.48); Probable helicase].
OS Strawberry mild yellow edge-associated virus (SMYEAV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.
OX NCBI_TaxID=12187;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MY-18;
RX MEDLINE=92166762; PubMed=1339469;
RA Jelkmann W., Maiss B., Martin R.R.:
RT "The nucleotide sequence and genome organization of strawberry mild
RT yellow edge-associated potexvirus.";
RL J. Gen. Virol. 73:475-479(1992).

-1- FUNCTION: RNA-replication. The central part of this protein
possibly functions as an ATP-binding helicase.
-1- CATALYTIC ACTIVITY: N nucleoside triphosphate +
(RNA)(N).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL: D12517; BAA02082.1; -;
CC PIR: J01426; RRMGSM.

DR InterPro: IPR001788; RNA_dep_RNAPol2.
DR InterPro: IPR000606; Viral_helicase1.
DR Pfam: PF00978; RNA_dep_RNAPol2_1.
DR Pfam: PF01443; Viral_helicase1_1.
KW ATP-binding; Helicase; RNA replication; RNA-directed RNA polymerase;
KW Transferrase.
FT NP_BIND 605 612 POTENTIAL.
SQ SEQUENCE 1323 AA; 149593 MW; D6BC35133E3CAD26 CRC64;

Query Match 69.0%; Score 40; DB 1; Length 1323;
Best Local Similarity 60.0%; Pred. No. 73;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 LOEWLKKKK 10
Db 615 LOEWMLSLKK 624

RESULT 4
ID PIG1_MOUSE STANDARD: PRT: 207 AA.
AC Q62077;

DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 1-phosphatidylinositol-4,5-bisphosphate phospholipase gamma 1
DE (EC 3.1.4.11) (PLC-gamma-1) (phospholipase C-gamma-1) (PLC-II)
DE (PLC-148) (Fragmant).
GN PLCG1 OR PLCG-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RX MEDLINE=96257754; PubMed=8687404;
RA Dupont G., McGuinness O.M., Johnson M.H., Berridge M.J.,
RA Borgese F.:
RT "Phospholipase C in mouse oocytes: characterization of beta and gamma
RT isoforms and their possible involvement in sperm-induced Ca2+
RT spiking.";
RL Biochem. J. 316:583-591(1996).

-1- FUNCTION: THE PRODUCTION OF THE SECOND MESSENGER MOLECULES
DIACYLGLYCEROL (DAG) AND INOSITOL 1,4,5-TRISPHOSPHATE (IP3) IS
MEDIATED BY ACTIVATED PHOSPHATIDYLINOSITOL-SPECIFIC PHOSPHOLIPASE
C ENZYMES.

-1- CATALYTIC ACTIVITY: 1-phosphatidy1-ID-myo-inositol 4,5-
bisphosphate + H(2)O -> D-myo-inositol 1,4,5-trisphosphate +
diacylglycerol.

-1- COFACTOR: Calcium.
-1- PTM: THE RECEPTOR-MEDIATED ACTIVATION OF PLC-GAMMA 1 AND 2
INVOLVES THEIR PHOSPHORYLATION BY TYROSINE KINASES IN RESPONSE
TO LIGATION OF A VARIETY OF GROWTH FACTOR RECEPTORS AND IMMUNE
SYSTEM RECEPTORS.

-1- SIMILARITY: DOMAINS X AND Y ARE CONSERVED IN DIFFERENT FORMS
OF PLC AND ARE ESSENTIAL FOR CATALYTIC ACTIVITY.
-1- SIMILARITY: CONTAINS 2 PH DOMAINS. THE SECOND ONE IS IN TWO
PARTS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL: X95346; CAA64639.1; -;
CC HSSP: P10688; IOAS.
CC MGD: MGI:97615; PLCG1.
CC InterPro: IPR000008; C2.

DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR000909; PL_PLC_Xdom.
 DR InterPro: IPR001711; PL_PLC_Y.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00169; PH; 1.
 DR Pfam: PF00387; PL_PLC-Y; 1.
 DR Prodom: PD001202; PL_PLC_Y; 1.
 DR SMART: SM00149; PLCYC; 1.
 DR PROSITE: PS00018; EF_HAND; PARTIAL.
 DR PROSITE: PS50002; SH3; PARTIAL.
 DR PROSITE: PS50003; PH_DOMAIN; 1.
 DR PROSITE: PS50004; C2_DOMAIN_2; PARTIAL.
 DR PROSITE: PS50007; PIPIC_X_DOMAIN; PARTIAL.
 DR PROSITE: PS50008; PIPIC_Y_DOMAIN; 1.
 KW Hydrolase: lipid degradation; Transducer; SH3 domain;
 KW Calcium-binding; Phosphorylation.
 FT NON_TER 1 1
 FT DOMAIN <1 6 SH3.
 FT DOMAIN 50 86 PH 2 (SECOND PART).
 FT DOMAIN 108 >207 DOMAIN Y.
 FT NON_TER 207 207
 SQ SEQUENCE 207 AA: 23750 MW: BBE228721381A61D CRC64:

Query Match 67.2%; Score 39; DB 1; Length 207;
 Best Local Similarity 45.5%; Pred. No. 16;
 Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOEWLKKLKKM 11
 Db 74 LQDWKKTREV 84

RESULT 5

VZ81_ARATH STANDARD: PRT: 209 AA.
 AC Q9S9T7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE VPS28 protein homolog 1.
 GN ATAG05000 OR T32M4.6.
 OS Arabidopsis thaliana (mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
 RA Harris B., Ansoyge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermayer B., Meche R., Mueller M.,
 RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
 RA Reichert B., Portetlelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohenseil J., Zimmermann W., Medler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,
 RA Breken M., Melkens I., Voet M., Bastiaens I., Aert R., Deftoor E.,
 RA Weltzenegger T., Bothe G., Ransperger U., Hilbert H., Braun M.,
 RA Holzner E., Brandt A., Peters S., van Staveren M., Dirkse W.,
 RA Moolman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,
 RA Petrelet A., Rajendram M.A., Lyne M., Benes V., Rechmann S.,
 RA Borokva D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fatlman B., Grandtath K., Danner D., Herzl A.,
 RA Neumann S., Argritov A., Vitale D., Lignori R., Piravandi E.,

RA Massenet O., Quigley F., Clabaud G., Muenzlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lechery A., Aubourg S.,
 RA Chedford F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
 RA Perez-Perez A., Fumelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedbia N., Gnoj L., Schultz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Thieleh J.,
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Mix P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Krimer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Gaisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekhar M., Mataro A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,
 RA Chen E., Marra M., Marliessen R., McCombie W.R.;
 RA "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana." 769-777 (1999).
 RL Nature 402:769-777 (1999).
 CC -1- SIMILARITY: BELONGS TO THE VPS28 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF162444; AAD48972.1; -
 DR EMBL: AL161502; CAB81042.1; -
 KW Hypothetical protein; Transport; Protein transport.
 SQ SEQUENCE 209 AA: 23558 MW: E4F226F1B91363E0 CRC64:

Query Match 67.2%; Score 39; DB 1; Length 209;
 Best Local Similarity 54.5%; Pred. No. 17;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKKLKKM 11
 Db 165 MKEWLSRISKM 175

RESULT 6

DPOL_SALTY STANDARD: PRT: 928 AA.
 AC Q9FL73;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase I (EC 2.7.7.7) (POL I).
 GN POLA OR SYM3999.
 OS Salmonella typhimurium.
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Salmonella.
 OX NCBI_TaxID=602;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2;
 RA Huang Y.P., Ito J.;
 RA "DNA polymerase I sequence from Salmonella typhimurium.";
 RA submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SCS1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Portwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,

RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterson R., Wilson R.K.:
 RT *Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 RT LT2.";
 RL Nature 413:852-856(2001).
 CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
 CC EXHIBITS 3' TO 5' AND 5' TO 3' EXONUCLEASE ACTIVITY. IT IS ABLE
 CC TO UTILIZE NICKED CIRCULAR DUPLEX DNA AS A TEMPLATE AND CAN
 CC UNWIND THE PARENTAL DNA STRAND FROM ITS TEMPLATE.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
 CC + (DNA)(N).
 CC -1- SUBUNIT: SINGLE-CHAIN MONOMER WITH MULTIPLE FUNCTIONS.
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch)
 CC -----
 DR EMBL: AF071212; AAG3170.1; -;
 DR EMBL: AE008886; AAL22838.1; -;
 DR HSSP: P00582; 1KLN.
 DR StyGene: SG77777; POLA.
 DR InterPro: IPR002562; 3_5_exonuclease.
 DR InterPro: IPR002421; 5_3_exonuclease.
 DR InterPro: IPR001098; DNA_pol.
 DR InterPro: IPR002298; DNA_pol.
 DR InterPro: IPR000513; Exo_N.I.
 DR InterPro: IPR003584; HNH_2.
 DR Pfam: PF00476; DNA_pol_A.1.
 DR Pfam: PF01367; 5_3_exonuclease.1.
 DR Pfam: PF01612; 3_5_exonuclease.1.
 DR Pfam: PF02739; 5_3_exonuc.N.1.
 DR SMART: SM00474; 3EXOC.1.
 DR SMART: SM00475; 53EXOC.1.
 DR SMART: SM00279; HNH2.1.
 DR SMART: SM00482; POLA.1.
 DR TIGRFAMS: TIGR00593; pola.1.
 DR PROSITE: PS00447; DNA_POLYMERASE_A.1.
 DR Transfaser: DNA-directed DNA polymerase; DNA replication; DNA repair;
 DR HydroLase; Nuclease; Exonuclease; DNA-binding; Complete proteome.
 FT DOMAIN 1 323 5'-3' EXONUCLEASE.
 FT DOMAIN 324 517 3'-5' EXONUCLEASE.
 FT DOMAIN 521 928 POLYMERASE.
 FT DOMAIN 324 928 KLENOW FRAGMENT.
 FT CONFLICT 751 751 G -> R (IN REF. 1).
 FT CONFLICT 848 848 M -> T (IN REF. 1).
 FT CONFLICT 867 867 L -> I (IN REF. 1).
 SQ SEQUENCE 928 AA: 103129 MW: FR887BF1039C160A CRC64;
 Query Match 67.2%; Score 39; DB 1; Length 928;
 Best Local Similarity 60.0%; Pred. No. 74;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LOEWLKLKK 10
 DB 339 LESWLEKLKK 348
 RESULT 7
 ID PIG1_HUMAN STANDARD; PRT; 1290 AA.
 AC P19174;
 DT 01-NOV-1990 (rel. 16, Created)
 DT 01-NOV-1990 (rel. 16, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)
 DE 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase gamma 1
 DE (EC 3.1.4.11) (PLC-gamma-1) (Phospholipase C-gamma-1) (PLC-II)
 DE (PLC-148).
 CC

GN PLCG1 OR PLC1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OC NCBI_TaxId=9606;
 RX SEQUENCE FROM N.A.
 RP MEDLINE-21638749; PubMed-11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grahnam D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Humble E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Levasalho M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McElay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehla H.K., Shownkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Walli M., Wallis J.M., Williams S.A.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Dubin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT *The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RP STRUCTURE BY NMR OF SH3 DOMAIN
 RX MEDLINE-93208890; PubMed-7681365;
 RA Kohda D., Hatanaka H., Oda M., Mandiyan V., Ullrich A.,
 RA Schlössinger J., Inagaki F.;
 RT *Solution structure of the SH3 domain of phospholipase C-gamma.";
 RL Cell 72:953-960(1993).
 CC -1- FUNCTION: PLC-GAMMA IS A MAJOR SUBSTRATE FOR HEPARIN-BINDING
 CC GROWTH FACTOR 1 (ACIDIC FIBROBLAST GROWTH FACTOR)-ACTIVATED
 CC TYROSINE KINASE.
 CC -1- CATALYTIC ACTIVITY: 1-phosphatidy-1D-myo-inositol 4,5-
 CC bisphosphate + H(2)O -> D-myo-inositol 1,4,5-trisphosphate +
 CC diacylglycerol.
 CC -1- COFACTOR: Calcium.
 CC -1- PTM: THE RECEPTOR-MEDIATED ACTIVATION OF PLC-GAMMA 1 AND 2
 CC INVOLVES THEIR PHOSPHORYLATION BY TYROSINE KINASES IN RESPONSE
 CC TO LIGATION OF A VARIETY OF GROWTH FACTOR RECEPTORS AND IMMUNE
 CC SYSTEM RECEPTORS.
 CC -1- SIMILARITY: DOMAINS X AND Y ARE CONSERVED IN DIFFERENT FORMS
 CC OF PLC AND ARE ESSENTIAL FOR CATALYTIC ACTIVITY.
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 PH DOMAINS. THE SECOND ONE IS IN TWO
 CC PARTS.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EF-HAND CALCIUM-BINDING DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M34667; AAA36452.1; -
 DR EMBL: AL022394; CAA18537.1; -
 DR PIR: A36466; A36466
 DR PDB: 2HSP; 31-AUG-94.
 DR PDB: 1HSQ; 31-AUG-94.
 DR Genew: HGNC:9065; PLCG1.
 DR MIM: 172420; -
 DR InterPro: IPR000008; C2.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR001192; PL-PLC.
 DR InterPro: IPR000909; PI-PLC_xdom.
 DR InterPro: IPR001711; PI-PLC_Y.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00168; C2; 1.
 DR Pfam: PF00169; PH; 2.
 DR Pfam: PF00387; PI-PLC-Y; 1.
 DR Pfam: PF00388; PI-PLC-X; 1.
 DR PRINTS: PR00390; PHPLIPASEC.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 2.
 DR ProDom: PD001202; PI-PLC_Y; 1.
 DR SMART: SM00239; C2; 1.
 DR SMART: SM00233; PH; 2.
 DR SMART: SM00146; PLCYC; 1.
 DR SMART: SM00149; PLCYC; 1.
 DR SMART: SM00252; SH2; 2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS00018; EF_HAND; 1.
 DR PROSITE: PS00001; SH2; 2.
 DR PROSITE: PS00002; SH3; 1.
 DR PROSITE: PS50003; PH_DOMAIN; 2.
 DR PROSITE: PS50004; C2_DOMAIN_2; 1.
 DR PROSITE: PS50007; PIPLC_X_DOMAIN; 1.
 DR PROSITE: PS50008; PIPLC_Y_DOMAIN; 1.
 KM Hydrolase; Lipid degradation; Transducer; SH2 domain; SH3 domain;
 Repeat; Calcium-binding; Phosphorylation; 3D-structure; Polymorphism.
 FT DOMAIN 27 142 PH 1.
 FT CA_BIND 165 176 EF_HAND (POTENTIAL).
 FT DOMAIN 320 464 DOMAIN X.
 FT DOMAIN 489 523 PH 2 (FIRST PART).
 FT DOMAIN 550 657 SH2 1.
 FT DOMAIN 668 756 SH2 2.
 FT DOMAIN 791 851 SH3.
 FT DOMAIN 895 931 PH 2 (SECOND PART).
 FT DOMAIN 953 1070 DOMAIN Y.
 FT DOMAIN 1075 1177 C2 DOMAIN.
 FT ACT_SITE 335 335 BY SIMILARITY.
 FT ACT_SITE 380 380 BY SIMILARITY.
 FT MOD_RES 771 771 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 783 783 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 1253 1253 PHOSPHORYLATION (BY SIMILARITY).
 FT VARIANT 813 813 I -> T (IN DBSNP:55381).
 FT SEQUENCE 1290 AA; 148531 MW; AE05ABE2A1BEDAC CRC64;
 /FTID=VAR_011908.

Query Match 67.2%; Score 39; DB 1; Length 1290;
 Best Local Similarity 45.5%; Pred. No. 1e+02;
 Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1 LQEWKLLKKM 11

Db 919 LQDWKXIREV 929

||||:||||:

RESULT 8

PGI1_RAT PGI1_RAT STANDARD; PRT; 1290 AA.

AC P10686;

DT 01-JUL-1989 (Rel. 11, Created)

DT 01-JUL-1989 (Rel. 11, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase gamma 1

DE (EC 3.1.4.11) (PLC-gamma-1) (phospholipase C-gamma-1) (PLC-II)

DE (PLC-148).

GN PIGL1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxId=10116;

RN [1]

RP SEQUENCE FROM N. A.

RA MEDLINE=88289733; PubMed=2840660;

RA Sun P.-G., Ryu S.H., Moon K.H., Suh H.W., Rhee S.G.;

RT "Inositol phospholipid-specific phospholipase C: complete cDNA and

RT protein sequences and sequence homology to tyrosine kinase-related

RT oncogene products.";

RL Proc. Natl. Acad. Sci. U.S.A. 85:5419-5423(1988).

CC -!- FUNCTION: THE PRODUCTION OF THE SECOND MESSENGER MOLECULES:

CC DIBUTYLGLYCEROL (DAG) AND INOSITOL 1,4,5-TRISPHOSPHATE (IP3) IS

CC MEDIATED BY ACTIVATED PHOSPHATIDYLINOSITOL-SPECIFIC PHOSPHOLIPASE

CC C ENZYMES.

CC -!- CATALYTIC ACTIVITY: 1-phosphatidy1-D-myo-inositol 4,5-

CC bisphosphate + H(2)O = D-myo-inositol 1,4,5-trisphosphate +

CC diacylglycerol.

CC -!- COFACTOR: Calcium.

CC -!- PTM: THE RECEPTOR-MEDIATED ACTIVATION OF PLC-GAMMA 1 AND 2

CC INVOLVES THEIR PHOSPHORYLATION BY TYROSINE KINASES IN RESPONSE

CC TO LIGATION OF A VARIETY OF GROWTH FACTOR RECEPTORS AND IMMUNE

CC SYSTEM RECEPTORS.

CC -!- SIMILARITY: DOMAINS X AND Y ARE CONSERVED IN DIFFERENT FORMS

CC OF PLC AND ARE ESSENTIAL FOR CATALYTIC ACTIVITY.

CC -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.

CC -!- SIMILARITY: CONTAINS 2 PH DOMAINS. THE SECOND ONE IS IN TWO

CC PARTS.

CC -!- SIMILARITY: CONTAINS 2 SH2 DOMAINS.

CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

CC -!- SIMILARITY: CONTAINS 1 EF-HAND CALCIUM-BINDING DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: J03806; AAA41921.1; -

DR PIR: A31317; A31317.

DR HSSP: P08487; 2PBD.

DR InterPro: IPR000008; C2.

DR InterPro: IPR002048; EF-hand.

DR InterPro: IPR001849; PH.

DR InterPro: IPR001192; PI-PLC.

DR InterPro: IPR000909; PI-PLC_xdom.

DR InterPro: IPR001711; PI-PLC_Y.

DR InterPro: IPR000980; SH2.

DR InterPro: IPR001452; SH3.

DR Pfam: PF00017; SH2; 2.

DR Pfam: PF00018; SH3; 1.

DR Pfam: PF00168; C2; 1.

DR Pfam: PF00169; PH; 2.

DR Pfam: PF00387; PI-PLC-Y; 1.

DR Pfam: PF00388; PI-PLC-X; 1.

DR PRINTS: PR00390; PPHPLIPASEC.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRINTS: PR00452; SH3DOMAIN.
DR Prodom: PD000066; SH3; 1.
DR Prodom: PD000093; SH2; 2.
DR Prodom: PD001202; PL_PLC_Y; 1.
DR SMART: SM00239; C2; 1.
DR SMART: SM00233; PH; 2.
DR SMART: SM00148; PLCYC; 1.
DR SMART: SM00149; PLCYC; 1.
DR SMART: SM00252; SH2; 2.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS00018; EF_HAND; 1.
DR PROSITE: PS50001; SH2; 2.
DR PROSITE: PS50002; SH3; 1.
DR PROSITE: PS50003; PH_DOMAIN; 2.
DR PROSITE: PS50004; C2_DOMAIN; 1.
DR PROSITE: PS50007; PIP2C_X_DOMAIN; 1.
DR PROSITE: PS50008; PIP2C_Y_DOMAIN; 1.
KW Hydrolase; Lipid degradation; Transducer; SH2 domain; SH3 domain;
KW Repeat; Calcium-binding; Phosphorylation.
FT CA_BIND 27 142 EF_HAND (POTENTIAL).
FT DOMAIN 320 464 DOMAIN X.
FT DOMAIN 489 523 PH 2 (FIRST PART).
FT DOMAIN 550 657 SH2 1.
FT DOMAIN 668 756 SH2 2.
FT DOMAIN 791 851 SH3.
FT DOMAIN 895 931 PH 2 (SECOND PART).
FT DOMAIN 953 1070 DOMAIN Y.
FT DOMAIN 1075 1177 C2 DOMAIN.
FT ACT_SITE 335 335 BY SIMILARITY.
FT ACT_SITE 380 380 BY SIMILARITY.
FT MOD_RES 771 771 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 783 783 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 1253 1253 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 1290 AA; 148547 MW; BB324DC27972CB3B CRC64;
Query Match 67.2%; Score 39; DB 1; Length 1290;
Best Local Similarity 45.5%; Pred. No. 1e+02;
Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
OY 1 LOEWLIKLIKMM 11
DB 919 LQDWVKRIREV 929
RESULT 9
PIGI_BOVIN STANDARD; PRT: 1291 AA.
AC P08487;
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-AUG-1988 (Rel. 41, Last annotation update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 1-phosphatidylinositol-4,5-bisphosphate phospholipase gamma 1
DE (EC 3.1.4.11) (PLC-gamma-1) (Phospholipase C-gamma-1) (PLC-II)
DE (PLC-148).
GN PLCG1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN 11
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=88156963; PubMed=2831461;
RA Stahl M.L., Ferenz C.R., Kelleher K.L., Kriz R.W., Knopf J.L.;
RT "Sequence similarity of phospholipase C with the non-catalytic region
of src.";
RL Nature 332:269-272(1988).
RN 12
RP PHOSPHORYLATION SITES.
RX MEDLINE=90154080; PubMed=1689310;

RA Kim J.W., Sim S.S., Kim U.H., Nishibe S., Whal M.I., Carpenter G.,
RA Rhee S.G.;
RT "Tyrosine residues in bovine phospholipase C-gamma phosphorylated by
RT the epidermal growth factor receptor in vitro.";
RL J. Biol. Chem. 265:3940-3943(1990).
RN [3]
RP PHOSPHORYLATION SITES.
RX MEDLINE=90154081; PubMed=1689311;
RA Whal M.I., Nishibe S., Kim J.W., Kim H.K., Rhee S.G., Carpenter G.;
RT "Identification of two epidermal growth factor-sensitive tyrosine
RT phosphorylation sites of phospholipase C-gamma in intact HSC-1
RT cells.";
RL J. Biol. Chem. 265:3944-3948(1990).
RN [4]
RP PHOSPHORYLATION SITES.
RX MEDLINE=91208680; PubMed=1708307;
RA Kim H.K., Kim J.W., Zilberstein A., Margolis B., Kim J.G.,
RA Schlessinger J., Rhee S.G.;
RT "PDGF stimulation of inositol phospholipid hydrolysis requires
RT PLC-gamma 1 phosphorylation on tyrosine residues 783 and 1254.";
RL Cell 65:435-441(1991).
RN [5]
RP STRUCTURE BY NMR OF 663-759.
RX MEDLINE=94236690; PubMed=8181064;
RA Pascal S.M., Singer A.U., Gish G., Yamazaki T., Shoelson S.E.,
RA Pawson T., Kay L.E., Forman-Kay J.D.;
RT "Nuclear magnetic resonance structure of an SH2 domain of
RT phospholipase C-gamma 1 complexed with a high affinity binding
RT peptide.";
RL Cell 77:461-472(1994).
CC -1- FUNCTION: THE PRODUCTION OF THE SECOND MESSENGER MOLECULES
CC DICYCLIGLUCEROL (DAG) AND INOSITOL 1,4,5-TRISPHOSPHATE (IP3) IS
CC MEDIATED BY ACTIVATED PHOSPHATIDYLINOSITOL-SPECIFIC PHOSPHOLIPASE
CC C ENZYMES.
CC -1- CATALYTIC ACTIVITY: 1-phosphatidy]-D-myo-inositol 4,5-
CC bisphosphate + H(2)O -> D-myo-inositol 1,4,5-trisphosphate +
CC diacylglycerol.
CC -1- COFACTOR: Calcium.
CC -1- PFM: THE RECEPTOR-MEDIATED ACTIVATION OF PLC-GAMMA 1 AND 2
CC INVOLVES THEIR PHOSPHORYLATION BY TYROSINE KINASES IN RESPONSE
CC TO LIGATION OF A VARIETY OF GROWTH FACTOR RECEPTORS AND IMMUNE
CC SYSTEM RECEPTORS.
CC -1- SIMILARITY: DOMAINS X AND Y ARE CONSERVED IN DIFFERENT FORMS
CC OF PLC AND ARE ESSENTIAL FOR CATALYTIC ACTIVITY.
CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 PH DOMAINS. THE SECOND ONE IS IN TWO
CC PARTS.
CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EF-HAND CALCIUM-BINDING DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
CC or send an email to license@sdb-sdb.ch).

CC EMBL: Y00301; CAA68406.1;
CC PIR: S00666; S00666.
CC PDB: 2PLD; 26-JAN-95.
CC PDB: 2PLE; 26-JAN-95.
CC InterPro: IPR000008; C2.
CC InterPro: IPR002048; EF-hand.
CC InterPro: IPR001849; PH.
CC InterPro: IPR001192; PL_PLC.
CC InterPro: IPR000909; PL_PLC_Xdom.
CC InterPro: IPR001711; PL_PLC_Y.
CC InterPro: IPR000980; SH2.
CC InterPro: IPR001452; SH3.
CC Pfam: PF00017; SH2; 2.
CC Pfam: PF00018; SH3; 1.

```
DR Pfam: PF00168; C2; 1.
DR Pfam: PF00169; PH; 2.
DR Pfam: PF00387; PI-PLC-Y; 1.
DR Pfam: PF00388; PI-PLC-X; 1.
DR PRINTS: PR00390; PHPLIPASEC.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRINTS: PR00452; SH3DOMAIN.
DR Prodom: PD000066; SH3; 1.
DR Prodom: PD000093; SH2; 2.
DR Prodom: PD001202; PI-PLC-Y; 1.
DR SMART: SM00239; C2; 1.
DR SMART: SM00239; PH; 2.
DR SMART: SM00148; PLCX; 1.
DR SMART: SM00149; PLCY; 1.
DR SMART: SM00252; SH2; 2.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS00018; EF_HAND; 1.
DR PROSITE: PS00001; SH2; 2.
DR PROSITE: PS50002; SH3; 1.
DR PROSITE: PS50003; PH_DOMAIN; 2.
DR PROSITE: PS50004; C2_DOMAIN; 2.
DR PROSITE: PS50007; PIPLC_X_DOMAIN; 1.
DR PROSITE: PS50008; PIPLC_Y_DOMAIN; 1.
KW Hydrolase; Lipid degradation; Transducer; SH2 domain; SH3 domain;
KW Repeat; Calcium-binding; Phosphorylation; 3D-structure.
FT DOMAIN 27 142 EF_HAND (POTENTIAL).
FT CA_BIND 165 176 DOMAIN X.
FT DOMAIN 320 464 PH 2 (FIRST PART).
FT DOMAIN 489 523 SH2 1.
FT DOMAIN 550 657 SH2 2.
FT DOMAIN 668 756 SH3.
FT DOMAIN 791 851 PH 2 (SECOND PART).
FT DOMAIN 895 931 DOMAIN Y.
FT DOMAIN 953 1070 C2 DOMAIN.
FT DOMAIN 1075 1177 BY SIMILARITY.
FT ACT_SITE 335 335 BY SIMILARITY.
FT ACT_SITE 380 380 BY SIMILARITY.
FT MOD_RES 771 771 PHOSPHORYLATION.
FT MOD_RES 783 783 PHOSPHORYLATION.
FT MOD_RES 1254 1254 PHOSPHORYLATION.
SO SEQUENCE 1291 AA; 148312 MW; 9F31C7DAA3F8E77 CRC64;

Query Match 67.2%; Score 39; DB 1; Length 1291;
Best Local Similarity 45.5%; Pred. No. 1e+02;
Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1 LQEWLKLK 11
Db 919 LQDWVKIKREV 929

RESULT 10
ACYO_CHICK STANDARD: PRT; 98 AA.
ID ACYO_CHICK STANDARD: PRT; 98 AA.
AC P07032.
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Acylphosphatase, organ-common type isozyme (EC 3.6.1.7)
DE (Acylphosphate phosphohydrolase) (Isozyme CH2).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RC TISSUE=Muscle;
RA MEDLINE=88139274; PubMed=2830254;
RA Ohba Y., Minowa O., Mizuno Y., Shiohawa H.;
RT "The primary structure of chicken muscle acylphosphatase isozyme
RT CH2."
RL J. Biochem. 102:1221-1229(1987).
```

```
CC -1- FUNCTION: ITS PHYSIOLOGICAL ROLE IS NOT YET CLEAR.
CC -1- CATALYTIC ACTIVITY: An acyl phosphate + H(2)O = a fatty acid anion
CC + phosphate
CC -1- TISSUE SPECIFICITY: ORGAN-COMMON TYPE ISOZYME IS FOUND IN MANY
CC DIFFERENT TISSUES.
CC -1- SIMILARITY: BELONGS TO THE ACYLPHOSPHATASE FAMILY.
DR PIR: A41513; OPCH2.
DR HSSP: P41500; ZACY.
DR InterPro: IPR001792; Acylphosphatase.
DR Pfam: PF00708; Acylphosphatase; 1.
DR PRINTS: PR00112; ACTLPHPTASE.
DR Prodom: PD001884; Acylphosphatase; 1.
DR PROSITE: PS00150; ACYLPHOSPHATASE_1; 1.
DR PROSITE: PS00151; ACYLPHOSPHATASE_2; 1.
KW Hydrolase; Acetylation; Multigene family.
FT MOD_RES 1 1 ACETYLATION.
FT ACT_SITE 23 1 POTENTIAL.
FT ACT_SITE 41 41 POTENTIAL.
SO SEQUENCE 98 AA; 11019 MW; 1AC35CF26218A7E5 CRC64;

Query Match 65.5%; Score 38; DB 1; Length 98;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 LQEWLKLK 8
Db 61 LQEWLRLK 68

RESULT 11
Y006_YEAST STANDARD: PRT; 313 AA.
ID Y006_YEAST STANDARD: PRT; 313 AA.
AC Q12094.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 35.7 kDa protein in DNL4-SIG1 intergenic region.
GN YOR006C OR UND313.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=97051599; PubMed=8896276;
RA Sterky F., Holmberg A., Pelterson B., Uhlen M.;
RT "The sequence of a 30 kb fragment on the left arm of chromosome XV
RT from Saccharomyces cerevisiae reveals 15 open reading frames, five of
RT which correspond to previously identified genes."
RL Yeast 12:1091-1095(1996).
CC -1- SIMILARITY: TO S. POMBE SPAC1F3.04C AND M. JANNASCHII MJ0708.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U43491; AAC49486.1; -
DR EMBL: Z74914; CAA99194.1; -
DR SGD: S0005532; YOR006C.
KW Hypothetical protein.
SO SEQUENCE 313 AA; 35686 MW; 6DF4703EA7252910 CRC64;

Query Match 65.5%; Score 38; DB 1; Length 313;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 QEWLKLK 10
Db 207 EEWLKLK 215
```


FT ACT_SITE 519 519 CATALYTIC BASE (BY SIMILARITY).
 FT ACT_SITE 522 522 GENERAL ACID CATALYST (BY SIMILARITY).
 FT ACT_SITE 523 523 INTERACT WITH SUBSTRATES (BY SIMILARITY).
 FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 323 323 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 415 415 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 424 424 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 514 514 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 547 547 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 646 646 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 651 651 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 721 721 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 742 742 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 164 164 MISSING (IN REF. 1).
 FT CONFLICT 624 624 D -> N (IN REF. 1).
 SQ SEQUENCE 768 AA; 82586 MW; 3FAC172C128A0C6F CRC64;

Query Match 65.5%; Score 38; DB 1; Length 768;
 Best Local Similarity 54.5%; Pred. No. 89;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKLKKM 11
 1:111:11:
 DB 339 LOEWLQROKV 349

RESULT 14
 YAI7_HUMAN STANDARD; PRT; 1015 AA.
 ID Q90P23;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE Hypothetical protein KIA1017.
 GN KIA1017.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=9246063; PubMed=10231032;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 6:63-70(1999).
 CC -1- SIMILARITY: TO HUMAN KIAA0297/KIAA0329.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AB023234; BAA76861.1; -;
 KM Hypothetical protein.
 SQ SEQUENCE 1015 AA; 114842 MW; 0214C2BFD67EA426 CRC64;

Query Match 65.5%; Score 38; DB 1; Length 1015;
 Best Local Similarity 70.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKLKK 10
 11111 1:1
 DB 526 LOEWLSLEK 535

RESULT 15
 YZ39_AQUAE STANDARD; PRT; 231 AA.
 ID YZ39_AQUAE
 AC O66427;
 DT 30-MAY-2000 (Rel. 39; Created)
 DT 30-MAY-2000 (Rel. 39; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE Hypothetical protein AA39.
 GN AA39.
 OS Aquifex aeolicus.
 OG Bacteria; Aquificae; Aquificales; Aquificaceae;
 OC Bacteria; Aquificae; Aquificales; Aquificaceae;
 OC Aquifex.
 OX NCBI_Taxid=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Snead M.A., Keller M., Anjay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus.";
 RL Nature 392:353-358(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE000667; AAC07979.1; -;
 KM Hypothetical protein; Plasmid; Complete proteome.
 SQ SEQUENCE 231 AA; 27534 MW; 162CFE422465C583 CRC64;

OY 1 LOEWLKLKK 10
 11111 1:1
 DB 54 LOEWLREKMK 64

RESULT 16
 RS16_CAUCR STANDARD; PRT; 165 AA.
 ID RS16_CAUCR
 AC P58122;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE 30S ribosomal protein S16.
 GN RPS16 OR CC3652.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_Taxid=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;

RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA Debey R.T., Dodson R.J., Durkin A.S., Gwyn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Utterback T., Tran K., Wolf A., Yamahyan J., Ernolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).

CC -1- SIMILARITY: BELONGS TO THE S16P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE006023; AAK25614.1; -
CC HSSP: P80379; 1EMW.
CC TIGR: CC3652; -
CC InterPro: IPR000307; Ribosomal_S16.
CC Pfam: PF00886; Ribosomal_S16; 1.
CC ProDom: PD003791; Ribosomal_S16; 1.
CC TIGRFAMs: TIGR00002; S16; 1.
CC PROSITE: PS00732; RIBOSOMAL_S16; 1.
CC Ribosomal protein: Complete proteome.
CC SEQUENCE 165 AA; 17605 MW; ED46FC2798C5B61C CRC64;
SQ
Query Match 63.8%; Score 37; DB 1; Length 165;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 QEWLTKK 7
DB 58 IQEWLTKK 64
RESULT 17
ID Y531_METUA STANDARD; PRT; 170 AA.
AC 057951;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M0531.
GN M0531.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=66880877;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uutterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RA jannaschii.";
RT Science 273:1058-1073(1996).
RL -1- SIMILARITY: BELONGS TO THE UPF0022 (USPA) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67502; AAB98522.1; -
CC HSSP: 057997; 1MJH.
CC TIGR: M0531;
CC InterPro: IPR000041; Usp.
CC Pfam: PF00582; Usp; 1.

DR PRINTS: PR01438; UNVRS1STRESS.
KW Hypothetical protein: Complete proteome
SQ SEQUENCE 170 AA; 18890 MW; D992BD530494964A CRC64;
Query Match 63.8%; Score 37; DB 1; Length 170;
Best Local Similarity 80.0%; Pred. No. 28;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 2 QEWLTKKKM 11
DB 89 QEWLTKKKM 98
RESULT 18
ID V282_ARATH STANDARD; PRT; 209 AA.
AC 065421;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE VPS28 protein homolog 2.
GN Atg621560 OR F18E5.180 OR F17L22.20.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entlan K.-D., Terryn N.,
RA Harris B., Anstorge W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermayer B., Mache R., Mueller M.,
RA Kleis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hohenseil J., Zimmermann W., Medler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Glymonprez B., Chuang Y.-C., Vandenbussche F.,
RA Breken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weltzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA Holzner E., Brandt A., Peters S., van Stevenen M., Dirkse W.,
RA Moolman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buysnaert C., Giejen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail A., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Leonard N., McIlroy K., Mayes R.,
RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefter M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fatlman B., Grandtath K., Dunner D., Herzl A.,
RA Neumann S., Argitluey A., Vitale D., Liguori R., Pflavandl E.,
RA Maassenet O., Quigley F., Clabaud G., Mendenhall A., Felder R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chedid F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
RA Perez-Perez A., Fumelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielek C.,
RA Fishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Hebermann K.,
RA Parnell L., Dedbia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Theieleh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Mix P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Krimer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Spletz J., Ryan E., Andrews S., Giesel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekhar M., Mateo A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,

```

RA Chen E., Marra M., Martienssen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RL thaliana";
CC
CC Nature 402:769-777(1999).
CC -1- SIMILARITY: BELONGS TO THE VPS28 FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: AL022603; CAA18720.1;
CC DR EMBL: AL035527; CAB36800.1;
CC DR EMBL: AL161555; CAB81263.1;
CC KW Hypothetical protein; transport; protein transport.
CC SEQUENCE 209 AA; 23494 MW; 8A8BBA415D9786B CRC64;
SO

Query Match 63.8%; Score 37; DB 1; Length 209;
Best Local Similarity 54.5%; Pred. No. 35;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LQEWLKLKLM 11
DB 165 MKEWLRLSKM 175

RESULT 19
BIOA_ECOLI STANDARD; PRT; 429 AA.
AC 123995;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Adenosylmethionine-8-amino-7-oxononanoate aminotransferase
DE (EC 2.6.1.62) (7,8-diamino-pelargonic acid aminotransferase) (DAPA
DE aminotransferase).
GN BIOA OR B0774.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
[1]
RX SEQUENCE FROM N.A.
RX MEDLINE=89066784; PubMed=3058702;
RA Otsuka A.J., Buocristian M.R., Howard P.K., Flamm J., Johnson O.,
RA Yamamoto R., Uchida K., Cook C., Ruppert J., Matsuzaki J.;
RT "The Escherichia coli biotin biosynthetic enzyme sequences predicted
RT from the nucleotide sequence of the bio operon.";
RL J. Biol. Chem. 263:19577-19585(1988).
RN [2]
RP SEQUENCE FROM N.A.
RP Pearson B.W., McKee R.A.;
RT "Genetic material for expression of biotin synthetase enzymes.";
RL Patent number GB2216530, 11-OCT-1989.
RN [3]
RP SEQUENCE FROM N.A.
RP STRAIN=K12 / MG1655;
RX MEDLINE=9742617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RX MEDLINE=99384134; PubMed=10452893;
RA Kack H., Sandmark J., Gibson K., Schneider G., Lindqvist Y.;
RT "Crystal structure of diaminopelargonic acid synthase: evolutionary
RT relationships between pyridoxal-5'-phosphate-dependent enzymes.";
```

```

RL J. Mol. Biol. 291:857-876(1999).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 8-amino-7-
CC oxononanoate -> S-adenosyl-4-methylthio-2-oxobutanoate + 7,8-
CC diaminononanoate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: Biotin biosynthesis.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: J04423; AAA23514.1;
CC DR EMBL: A11524; CAA00964.1;
CC DR EMBL: AE000180; AAC73861.1;
CC DR PIR: A32025; XNECDP.
CC DR PDB: 1QJ3; 22-JUN-00.
CC DR PDB: 1QJ5; 22-JUN-00.
CC DR EcoGene: EG10117; bioA.
CC DR InterPro: IPR000954; Aminotran_3.
CC DR Pfam: PF00202; aminotran_3.
CC DR TIGRfams: TIGR00508; bioA_1;
CC DR PROSITE: PS00600; AA-TRANSFER_CLASS_3; 1.
CC KW Biotin biosynthesis; Transferase; Aminotransferase;
CC Pyridoxal phosphate; Complete proteome; 3D-structure.
CC FT BINDING 274 274
CC FT CONFLICT 11 11
CC FT CONFLICT 99 102
CC FT SEQUENCE 429 AA; 47335 MW; 84D2D1AE3A1280FF CRC64;
SO

Query Match 63.8%; Score 37; DB 1; Length 429;
Best Local Similarity 55.6%; Pred. No. 72;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 3 EWLKKLKM 11
DB 226 EWLKIRKI 234

RESULT 20
GIGA_BACST STANDARD; PRT; 485 AA.
AC 008328;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glycogen synthase (EC 2.4.1.21) [Starch [bacterial glycogen]
DE synthase).
DE GIGA.
GN Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Geobacillus.
OX NCBI_TaxID=1422;
[1]
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=TBRE14;
RX MEDLINE=97386405; PubMed=9244254;
RA Takata H., Takata T., Okada S., Takagi M., Imanaka T.;
RT "Characterization of a gene cluster for glycogen biosynthesis and a
RT heterotrimeric ADP-glucose pyrophosphorylase from Bacillus
RT stearothermophilus.";
RL J. Bacteriol. 179:4689-4698(1997).
CC -1- FUNCTION: Synthesizes alpha-1,4-glucan chains using ADP-glucose.
CC -1- CATALYTIC ACTIVITY: ADP-glucose + ((1,4)-alpha-D-glucosyl)(N) ->
CC ADP + ((1,4)-alpha-D-glucosyl)(N+1).
CC -1- PATHWAY: Glycogen biosynthesis; second step.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
```



```

CC      FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: D87026; BAA19591.1; -
DR      InterPro: IPR001296; Glycos_transf_1.
DR      Pfam: PF00534; Glycos_transf_1.1.
KW      Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT      BINDING 15 15 ADP-GLUCOSE (BY SIMILARITY).
SQ      SEQUENCE 485 AA; 56244 MW; AEE338223ECCEAB CRC64;

Query Match 63.8%; Score 37; DB 1; Length 485;
Best Local Similarity 66.7%; Pred. No. 81;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY      3 EMLKLRKK 11
DB      50 EMLKLRKK 58

RESULT 21
TRBE_RHISN STANDARD: PRT: 662 AA.
AC      P53539;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      Probable conjugal transfer protein trbe part 2.
GN      TRBEA OR Y4CO.
OS      Rhizobium sp. (strain NGR234).
OC      Plasmid sym pNGR234a.
OC      Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC      Rhizobiaceae; Rhizobium.
OX      NCBI_TaxID=394;
RN      [1]
RP      MEDLINE:97305956; PubMed-9163424;
RA      Freiberg C.A., Fellay R., Broughton W.J., Rosenthal A.,
RA      Perret X.;
RT      "Molecular basis of symbiosis between Rhizobium and legumes.";
RL      Nature 387:394-401(1997).
CC      -1- SIMILARITY: STRONG, TO C-TERMINAL OF A.TUMEFACIENS TI PLASMID
CC      TRBE.
CC      -1- SIMILARITY: ALSO TO A.TUMEFACIENS VIRB4.
CC      -1- CAUTION: IT IS POSSIBLE THAT TRBEA AND TRBEA FORM A SINGLE ORF.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: AE000066; AAB92432.1; -
DR      InterPro: IPR003593; AAA_ATPase.
DR      InterPro: IPR004346; CAGE_TRBE_VIRB.
DR      Pfam: PF03135; CAGE_TRBE_VIRB; 1.
DR      SMART: SM00382; AAA; 1.
KW      Conjugation; Plasmid; ATP-binding.
FT      NP_BIND 307 314 ATP (POTENTIAL).
SQ      SEQUENCE 662 AA; 74393 MW; 2D0233063754D9C CRC64;

Query Match 63.8%; Score 37; DB 1; Length 662;
Best Local Similarity 60.0%; Pred. No. 1.le02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      1 LOEWLKLKK 10
DB      518 IREWLKVLRK 527

RESULT 22
TRBE_AGRU STANDARD: PRT: 820 AA.
AC      P54910;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Conjugal transfer protein trbe precursor.
GN      TRBE.
OS      Agrobacterium tumefaciens.
OC      Plasmid pTiA6NC.
OC      Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC      Rhizobiaceae; Rhizobium.
OX      NCBI_TaxID=358;
RN      [1]
RP      MEDLINE:96312368; PubMed-8763954;
RA      Alt-Morbe J., Stryker J.L., Fuqua C., Li P.L., Farrand S.K.,
RA      Winans S.C.;
RT      "The conjugal transfer system of Agrobacterium tumefaciens
RT      octopine-type Ti plasmids is closely related to Ti plasmid vir genes.";
RL      J. Bacteriol. 178:4248-4257(1996).
CC      -1- SIMILARITY: TO A.TUMEFACIENS VIRB4.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: AF242881; AAB95097.1; -
DR      InterPro: IPR003593; AAA_ATPase.
DR      InterPro: IPR004346; CAGE_TRBE_VIRB.
DR      Pfam: PF03135; CAGE_TRBE_VIRB; 1.
DR      SMART: SM00382; AAA; 1.
DR      TIGRFAMs: TIGR00929; VIRB4_CAGE; 1.
KW      Conjugation; Plasmid; ATP-binding; Signal.
FT      SIGNAL 1 15
FT      CHAIN 16 820 CONJUGAL TRANSFER PROTEIN TRBE.
FT      NP_BIND 463 470 ATP (POTENTIAL).
SQ      SEQUENCE 820 AA; 91536 MW; DEA3A2A84DE7A4E1 CRC64;

Query Match 63.8%; Score 37; DB 1; Length 820;
Best Local Similarity 60.0%; Pred. No. 1.4e02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY      1 LOEWLKLKK 10
DB      676 IREWLKVLRK 685

RESULT 23
MSH2_YEAST STANDARD: PRT: 964 AA.
AC      P25847; 012423;
DT      01-MAY-1997 (Rel. 22, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      DNA mismatch repair protein MSH2.
GN      MSH2 OR YOL090W OR O0935.
OS      Saccharomyces cerevisiae (Baker's yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX      NCBI_TaxID=4932;
RN      [1]

```

RP SEQUENCE FROM N.A.
 RX MEDLINE=93093456; PubMed=1459447;
 RA Reenan R.A.G., Kolodner R.D.;
 RT "Isolation and characterization of two Saccharomyces cerevisiae genes
 RT encoding homologs of the bacterial Hexa and Muts mismatch repair
 RT proteins.";
 RL Genetics 132:963-973(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1679;
 RX MEDLINE=96021609; PubMed=8533473;
 RA Zumbstein E., Pearson B.M., Kalogeropoulos A., Schweizer M.;
 RT "A 29.425 kb segment on the left arm of yeast chromosome XV contains
 RT more than twice as many unknown as known open reading frames.";
 RL yeast 11:975-986(1995).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=93093457; PubMed=1334021;
 RA Reenan R.A.G., Kolodner R.D.;
 RT "Characterization of insertion mutations in the Saccharomyces
 RT cerevisiae MSH1 and MSH2 genes: evidence for separate mitochondrial
 RT and nuclear functions.";
 RL Genetics 132:975-985(1992).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=96413311; PubMed=8816473;
 RA Alani E.;
 RT "The Saccharomyces cerevisiae Msh2 and Msh6 proteins form a complex
 RT that specifically binds to duplex oligonucleotides containing
 RT mismatched DNA base pairs.";
 RL Mol. Cell. Biol. 16:5604-5615(1996).
 RN [5]
 RP MUTAGENESIS OF GLY-317; LEU-402; GLN-430; ASP-524; ARG-542; PRO-640
 RP AND CYS-716.
 RX MEDLINE=99400995; PubMed=10469597;
 RA Droschmann K., Clark A.B., Kunkel T.A.;
 RT "Mutator phenotypes of common polymorphisms and missense mutations in
 RT MSH2.";
 RL Curr. Biol. 9:907-910(1999).
 RN [6]
 RP MUTAGENESIS OF GLY-317; PRO-640 AND HIS-658.
 RX MEDLINE=21439334; PubMed=11555625;
 RA Ellison A.R., Lofing J., Bitter G.A.;
 RT "Functional analysis of human MLH1 and MSH2 missense variants and
 RT hybrid human-yeast MLH1 proteins in Saccharomyces cerevisiae.";
 RL Hum. Mol. Genet. 10:1889-1900(2001).
 CC -1- FUNCTION: INVOLVED IN POST-REPLICATIVE DNA-MISMATCH REPAIR. BINDS
 CC TO MISMATCH-CONTAINING DNA.
 CC -1- SUBUNIT: HETERODIMER OF MSH2 AND MSH6.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M84170; AAA34802.1; -
 DR EMBL: X83121; CAA58189.1; -
 DR EMBL: Z74832; CAA99102.1; -
 DR PIR: S27433; S27433.
 DR SGD: S0005450; MSH2.
 DR InterPro: IPR000432; MUTS_C.
 DR InterPro: IPR002863; MUTS_N.
 DR Pfam: PF01624; MUTS_C; 1.
 DR Pfam: PF01624; MUTS_N; 1.
 DR ProDom: PD001263; MUTS_C; 1.
 DR SMART: SM00534; MUTSac; 1.
 DR SMART: SM00533; MUTSd; 1.
 DR PROSITE: PS00486; DNA_MISMATCH_REPAIR_2; 1.

KW DNA repair; ATP-binding; DNA-binding; Nuclear protein.
 FT NP_BIND 688 695 ATP (POTENTIAL).
 FT MUTAGEN 317 317 G->D: PARTIALLY DEFECTIVE IN A MISMATCH
 FT MUTAGEN 402 402 REPAIR ASSAY.
 FT MUTAGEN 402 402 L->F: PARTIALLY DEFECTIVE IN A MISMATCH
 FT MUTAGEN 430 430 REPAIR ASSAY.
 FT MUTAGEN 430 430 O->K: PARTIALLY DEFECTIVE IN A MISMATCH
 FT MUTAGEN 524 524 REPAIR ASSAY.
 FT MUTAGEN 524 524 D->Y: PARTIALLY DEFECTIVE IN A MISMATCH
 FT MUTAGEN 542 542 REPAIR ASSAY.
 FT MUTAGEN 640 640 R->P: DEFECTIVE IN A MISMATCH REPAIR
 FT MUTAGEN 658 658 P->L: DEFECTIVE IN A MISMATCH REPAIR
 FT MUTAGEN 716 716 ASSAY.
 FT MUTAGEN 716 716 H->Y: FULLY FUNCTIONAL IN A MISMATCH
 FT MUTAGEN 716 716 REPAIR ASSAY.
 FT MUTAGEN 716 716 C->E: DEFECTIVE IN A MISMATCH REPAIR
 FT MUTAGEN 716 716 ASSAY.
 FT MUTAGEN 716 716 KYKALL -> EYKSPCCYN (IN REF. 1).
 SO SEQUENCE 964 AA; 108884 MW; 43FPD8A640138AE4 CRC64;
 Query Match 63.8%; Score 37; DB 1; Length 964;
 Best Local Similarity 50.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
 OY 1 LOEWLKLKK 10
 Db 910 LKEMIRKYKE 919
 RESULT 24
 EXSC_BUCAL STANDARD; PRT: 1070 AA.
 ID EXSC_BUCAL
 AC P57528;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Exodeoxyribonuclease V gamma chain (EC 3.1.11.5).
 GN RECC OR BU453.
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 OS symbiotic bacterium).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=118099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tokyo 1998;
 RX MEDLINE=20445173; PubMed=10993077;
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 RT Buchnera sp. APS.";
 RL Nature 407:81-86(2000).
 CC -1- FUNCTION: EXHIBITS A WIDE VARIETY OF CATALYTIC ACTIVITIES
 CC INCLUDING ATP-DEPENDENT EXONUCLEASE, ATP-STIMULATED ENDONUCLEASE,
 CC ATP-DEPENDENT HELICASE AND DNA-DEPENDENT ATPASE ACTIVITIES
 CC (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage (in the presence of
 CC ATP) in either 5'- to 3'- or 3'- to 5'-direction to yield 5'-
 CC phosphooligonucleotides.
 CC -1- SUBUNIT: CONSIST OF THREE SUBUNITS; RECB, RECC AND RECD
 CC (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AP001119; BAB13151.1; -
 DR HydroLase: Nuclease; Exonuclease; Endonuclease; Helicase; DNA repair;
 KW Complete proteome.
 SO SEQUENCE .1070 AA; 128578 MW; 8F644C84877981AF CRC64;

Query Match 63.8%; Score 37; DB 1; Length 1070;
 Best Local Similarity 54.5%; Pred. No. 1.8e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKKLKK 11
 DB 526 LEKWKKKLKK 536

RESULT 25

RPPO_NMV STANDARD; PRT: 1643 AA.

AC P15095;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE RNA replication protein (186 kDa protein) (ORF 1) [Contains: RNA-directed RNA polymerase (EC 2.7.7.48); Probable helicase].
 OS Narcissus moscia virus (NMV).
 OC Viruses: ssRNA positive-strand viruses, no DNA stage; Potexvirus.
 OX NCBI_TaxID=12180;

RP SEQUENCE FROM N.A.
 RX MEDLINE=89279206; PubMed=2732689;
 RA Zuidema D., Linthorst H.J.M., Huisman M.J., Asjes C.J., Bol J.F.;
 RT "Nucleotide sequence of narcissus moscia virus RNA."
 RL J. Gen. Virol. 70:267-276(1989).
 CC -1- FUNCTION: RNA-replication. The central part of this protein
 CC possibly functions as an ATP-binding helicase.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate +
 CC [RNA](N).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D13747; BAA02891.1;
 DR PIR: J70470; RRMGNV.
 DR InterPro: IPR001788; RNA_dep_RNApol2.
 DR InterPro: IPR000606; Viral_helicase1.
 DR Pfam: PF00978; RNA_dep_RNApol2; 1.
 DR Pfam: PF01443; Viral_helicase1; 1.
 KW ATP-binding; Helicase; RNA replication; RNA-directed RNA polymerase;
 KW Transferase.
 FT NP_BIND 868 875 ATP (POTENTIAL).
 SO SEQUENCE 1643 AA; 186304 MW; ADA66FC0FE2EC43 CRC64;

Query Match 63.8%; Score 37; DB 1; Length 1643;
 Best Local Similarity 50.0%; Pred. No. 2.8e+02;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKKLKK 10
 DB 878 IQEMWRSLLKE 887

RESULT 26
 XKDI_BACSU STANDARD; PRT: 161 AA.

DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phage-like element PBX protein xkdi.
 GN XKDI.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=168;
 RA Krogh S., O'Reilly M., Nolan N., Devine K.M.;
 RL Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=168;
 RA MEDLINE=98044033; PubMed=9384377;

RA Kunz F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabre C., Ferrati E., Fougere D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Maun C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogilawa A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porvolik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche E., Roche R., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Tarpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzinger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis".
 RL Nature 390:249-256(1997).

CC -1- SIMILARITY: STRONG, TO B.SUBTILIS Y081.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Z70177; CA94064.1;
 DR EMBL: Z99110; CAB13120.1;
 DR Subtilist; BG1544; xkdi.
 KW Complete proteome.
 SO SEQUENCE 161 AA; 18262 MW; C45CF616CD90BD16 CRC64;

Query Match 62.1%; Score 36; DB 1; Length 161;
 Best Local Similarity 75.0%; Pred. No. 39;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOEWLKKL 8
 DB 154 LRQWLKKL 161

RESULT 27

COAE_MYCPN STANDARD; PRT: 200 AA.

AC P75400;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Dephospho-CoA kinase (EC 2.7.1.24) (Dephosphocoenzyme A kinase).

```

GN COAE OR MPN382 OR MP455.
OS Mycoplasma pneumoniae.
OC Bacteria: Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
  SEQUENCE FROM N.A.
RC STRAIN-ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfrich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
  Hermann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
  pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- FUNCTION: CATALYZES THE PHOSPHORYLATION OF THE 3'-HYDROXYL GROUP
  OF DEPHOSPHOCOENZYME A TO FORM COENZYME A (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + dephospho-CoA = ADP + CoA.
CC -1- PATHWAY: Coenzyme A (CoA) biosynthesis; fifth (last) step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE COAE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AE000044; BAB96103.1; -
DR InterPro: IPR001977; Depp_COAKinase.
DR Pfam: PF01121; Coae; 1.
DR ProDom: PD003329; Depp_COAKinase; 1.
DR TIGRFAMs: TIGR00152; UPE0038; 1.
DR PROSITE: PS01294; COAE; FALSE NEG.
KM Transferase: Kinase; ATP-binding; Coenzyme A biosynthesis;
KW Complete proteome.
FT NP_BIND 7 14 ATP (POTENTIAL).
SQ SEQUENCE 200 AA; 22910 MW; 396EAD0C4AA940589 CRC64;

Query Match
Best Local Similarity 62.1%; Score 36; DB 1; Length 200;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKLK 9
DB 90 VOEWLNQLK 98

RESULT 28
RB21_CANFA STANDARD; PRT; 223 AA.
AC P55745;
DT 01-NOV-1997 (Rel. 35; Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DT 15-DEC-1998 (Rel. 37; Last annotation update)
DE Ras-related protein Rab-21.
GN RAB21.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Carnivora; Flsipedida; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
  SEQUENCE FROM N.A.
RA Chavrier P.;
RT "Rab21."
RL (in) Zerial M., Huber L.A. (eds.);
  Guidebook to the small GTPases, pp.365-366, Oxford University Press,
  Oxford (1995).
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC COMPARTMENTS AND PLASMA MEMBRANE.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR HSSP: P36017; IER0
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_transferring.

```

```

DR InterPro: IPR005225; Small_GTP.
DR Pfam: PR00071; ras; 1
DR PRINTS: PR00449; RASTRNSFRNG.
DR SMART: SM00175; Rab; 1.
DR TIGRFAMs: TIGR00231; small_GTP; 1.
KM GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 24 31 GTP (BY SIMILARITY).
FT NP_BIND 72 76 GTP (BY SIMILARITY).
FT NP_BIND 130 133 GTP (BY SIMILARITY).
FT DOMAIN 46 54 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 219 219 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 220 220 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 223 AA; 24161 MW; B9D198C48442294D CRC64;

Query Match
Best Local Similarity 45.5%; Score 36; DB 1; Length 223;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKLK 11
DB 108 VKNWVKELRKM 118

RESULT 29
RB21_HUMAN STANDARD; PRT; 225 AA.
AC Q9UJ25; Q14466;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Ras-related protein Rab-21.
GN RAB21 OR KIAA0118.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
  SEQUENCE FROM N.A.
RP TISSUE=Colon carcinoma;
RA Oudem F.J.M., van Bokhoven H., Kamps G., Ginsel L.A., Franssen J.A.M.;
RT "Rab GTPases expressed in Caco-2 cells."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
  SEQUENCE FROM N.A.
RP TISSUE=Lymph;
RA Strusberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
  SEQUENCE OF 65-225 FROM N.A.
RX MEDLINE=95308325; PubMed=7788527;
RA Nagase T., Miyajima N., Tanaka A., Sazuka T., Seki N., Sato S.,
  Tabata S., Ishikawa K.-I., Kawarabayashi Y., Kotani H., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. III.
  The coding sequences of 40 new genes (K14A0081-K14A0120) deduced by
  analysis of cDNA clones from human cell line KG-1."
RL DNA Res. 2:37-43(1995).
CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE GOLGI APPARATUS (BY
  SIMILARITY).
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF091035; AAP00048.1; -
DR EMBL: BC021901; AAH21901.1; -
DR EMBL: D42087; BAA07682.1; -
DR HSSP: P36017; IER0
DR Genew: HGNC:18263; RAB21.

```

```

DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_tnsifmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMNG.
DR SMART: SM00175; RAB; 1.
DR TIGRFAMs: TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 26 33 GTP (BY SIMILARITY).
FT NP_BIND 74 78 GTP (BY SIMILARITY).
FT NP_BIND 132 135 GTP (BY SIMILARITY).
FT DOMAIN 48 56 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 221 221 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 222 222 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 225 AA: 24347 MW: 730AC2C127FCBFD6 CRC64;

Query Match 62.1%; Score 36; DB 1; Length 225;
Best Local Similarity 45.5%; Pred. No. 54;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKLKKM 11
DB 110 VKNWVKELRKM 120

RESULT 30
2NPD_WILMR STANDARD: PRT: 374 AA.
AC 012723:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 2-nitropropane dioxygenase (EC 1.13.11.32) (Nitroalkane oxidase)
DE (2-NPD).
OS Williopsis mrakii (Yeast) (Hansenula mrakii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Williopsis.
OX NCBI_TaxID=4963;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-IFO 0895;
RX MEDLINE-95112848; PubMed-7813473;
RT Tchorzewski M., Kurinara T., Esaki N., Soda K.;
RT "Unique primary structure of 2-nitropropane dioxygenase from
RT Hansenula mrakii."
RL Eur. J. Biochem. 226:841-846(1994).
CC -1- FUNCTION: CATALYZES THE OXYGENATIVE DENITRIFICATION OF VARIOUS
CC ANIONIC NITROALKANES.
CC -1- CATALYTIC ACTIVITY: 2 2-nitropropane + O(2) = 2 acetone + 2
CC nitrite.
CC -1- COFACTOR: FAD.
CC -1- SIMILARITY: SOME. TO FMN-DEPENDENT 2-NITROPROPANE DIOXYGENASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U13900; AAA64484.1; -.
DR InterPro: IPR004136; 2npdp_dioxygen.
DR Pfam: PF03060; NPD; 1.
KW Oxidoreductase; Dioxygenase; Flavoprotein; FAD.
SQ SEQUENCE 374 AA: 41467 MW: 63f828b7453e3f0 CRC64;

Query Match 62.1%; Score 36; DB 1; Length 374;
Best Local Similarity 60.0%; Pred. No. 91;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 QEWLKLKKM 11
DB 110 VKNWVKELRKM 120

```

```

DB 93 EEWLKKYDKI 102
RESULT 31
SYN_ECOLI STANDARD: PRT: 423 AA.
AC P00951;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosyl-tRNA synthetase (EC 6.1.1.1) (Tyrosine--tRNA ligase) (TyrRS).
GN TYRS OR B137 OR Z2650 OR ECS2346.
OS Escherichia coli.
OS Escherichia coli 0157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562; 83334;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC MEDLINE-83132295; PubMed-6761148;
RX Barker D.G., Bruton C.J., Winter G.;
RT "The tyrosyl-tRNA synthetase from Escherichia coli. Complete
RT nucleotide sequence of the structural gene."
RL FEBS Lett. 150:419-423(1982).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE-97251357; PubMed-9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizouchi K., Mori H., Mori T.,
RA Motomura K., Nakado S., Nakamura Y., Nashimoto H., Nishio Y.,
RA Oshima T., Saito N., Sampei G., Seki Y., Sivasubram S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 mln region on the linkage map."
RL DNA Res. 3:363-377(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-0157:H7 / EDL933 / ATCC 700927;
RX MEDLINE-21074935; PubMed-11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dialanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blatner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli 0157:H7."
RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN-0157:H7 / RIMD 0509952;
RX MEDLINE-21156231; PubMed-11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT 0157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:111-224(2001).
RN [6]
RP SEQUENCE OF 1-25 FROM N.A.

```


Matches	5	Conservative	4	Mismatches	0	Indels	0	Gaps	0
OY	3	EWLKKTKRM 11							
	11:::111:								
Db	441	EWIRELKKV 449							
RESULT 33									
M5_STRPS									
ID	M5_STRPS	STANDARD:	PRT:	492 AA.					
AC	P02977;								
DT	21-JUL-1986 (Rel. 01, Created)								
DT	01-FEB-1991 (Rel. 17, Last sequence update)								
DT	15-JUN-2002 (Rel. 41, Last annotation update)								
DE	M protein, serotype 5 precursor.								
CN	EMM5 OR SMP5.								
OS	Streptococcus pyogenes (serotype M5).								
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;								
OC	Streptococcus.								
OX	NCBI_TaxID=160491;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RA	MEDLINE=88186881; PubMed=3281944;								
RA	Miller L., Gray L., Beachey E., Kehoe M.;								
RT	"Antigenic variation among group A streptococcal M proteins.								
RT	Nucleotide sequence of the serotype 5 M protein gene and its								
RT	relationship with genes encoding types 6 and 24 M proteins.";								
RL	J. Biol. Chem. 263:5668-5673(1988).								
RN	[2]								
RP	SEQUENCE OF 43-212 AND 238-250.								
RX	MEDLINE=84162039; PubMed=6368549;								
RA	Manjula B.N., Acharya A.S., Mische S.M., Fairwell T., Fischetti V.A.;								
RT	"The complete amino acid sequence of a biologically active								
RT	197-residue fragment of M protein isolated from type 5 group A								
RT	streptococci.";								
RL	J. Biol. Chem. 259:3686-3693(1984).								
CC	-1- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES								
CC	OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF								
CC	THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO								
CC	PHAGOCYTOSIS.								
CC	-1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by								
CC	an amide bond (Potential).								
CC	-1- SIMILARITY: TO OTHER M PROTEINS.								
CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation at								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage of and for commercial								
CC	entities requires a license agreement (see http://www.isb-sdb.ch/announce/								
CC	or send an email to license@sdb-sdb.ch).								
CC	-----								
DR	EMBL; M20374; AAA26976.1; -.								
DR	PIR; A03501; MMSOMP.								
DR	PIR; A28616; A28616.								
DR	InterPro: IPR001899; Gram_pos_anchor.								
DR	InterPro: IPR003345; M.repeat.								
DR	Pfam; PF00746; Gram_pos_anchor: 1.								
DR	Pfam; PF02370; M; 9.								
DR	PRINTS; PR00015; GPOSANCHOR.								
DR	TIGRFAMS; TIGR01167; LPXTG_anchor: 1.								
DR	TIGRFAMS; TIGR01168; YSIRK_signal: 1.								
DR	PROSITE: PS50847; GRAM_POS_ANCHORING: 1.								
KW	Virulence; Phagocytosis; Cell wall; Peptidoglycan-anchor; Repeat;								
KW	Antigen; Coiled coil; Signal.								
KW	SIGNAL	1	42						
FT	CHAIN	43	461						
FT	PROPEP	462	492						
FT	DOMAIN								

```

FT REPEAT 90 96 4.
FT SITE 458 462 5.
FT MOD_RES 461 461 4.
FT CONFLICT 43 43 4.
FT CONFLICT 50 50 5.
FT CONFLICT 102 102 5.
FT CONFLICT 208 208 5.
SO SEQUENCE 492 AA; 55085 MW; 873779B6C8D5E27 CRC64;

Query Match
Best Local Similarity 62.1%; Score 36; DB 1; Length 492;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 WLKRLKK 10
|||
Db 11 WLKRLKK 17

RESULT 34
VG28_BPMU STANDARD; PRT; 551 AA.
AC O9T1W6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein gp28.
GN Bacteriophage Mu.
OS Viruses: dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC Mu-like viruses;
CX NCBI_TaxID=10677;
RN [1]
RP SEQUENCE FROM N.A.
RA Morgan G., Hatfull G., Hendrix R.;
RT "Genome of bacteriophage Mu and comparison with the Haemophilus
influenzae mu-like prophage PluM."
RL Submitted (AUG-1998) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: PUTATIVE PORTAL PROTEIN.
CC -1- SIMILARITY: STRONG, TO H.INFLUENZAE HI1500.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF083977; AAF01106.1;
SQ SEQUENCE 551 AA; 62635 MW; B63D646144653A7C CRC64;

Query Match
Best Local Similarity 62.1%; Score 36; DB 1; Length 551;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 QEWLKKK 9
|||
Db 443 QEWMPKKK 450

RESULT 35
FBF2_CAEEL STANDARD; PRT; 632 AA.
AC 009312;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein Fbf-2.
GN FBF-2 OR F2H12.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabdilitoidea;
OC Rhabdilitidae; Pelodermidae; Caenorhabdilitis.

```

OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol NZ:
 RA Pavilio T.
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE PUMILIO/MPT5 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U23176; AAC46716.1; -
 DR Wormpep; F21H12.5; CE01916.
 DR InterPro; IPR001313; Pumilio/Puf.
 DR Pfam; PF00806; PUF; 7.
 DR SMART; SM00025; Pumilio; 8.
 DR Repeat.
 KW SEQUENCE 632 AA; 71805 MW; 6EDE86F71E4386 CRC64;
 SQ
 Query Match: 62.1%; Score 36; DB 1; Length 632;
 Best Local Similarity 85.7%; Pred. NO. 1.5e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 2 QEWLKKL 8
 DB 532 QDWLKKL 538
 RESULT 36
 PIXB_HUMAN STANDARD; PRT; 646 AA.
 ID Q14155;
 AC 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Rho guanine nucleotide exchange factor 7 (PAK-interacting exchange
 DE factor beta) (Beta-Pix) (COOL-1) (P85).
 GN ARHGEF7 OR PAK3BP OR PIXB OR COOL1 OR KIAA0142.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RX MEDLINE=96127530; PubMed=8590280;
 RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
 RT "Prediction of the coding sequences of unidentified human genes. IV.
 RT The coding sequences of 40 new genes (K1AA0121-K1AA0160) deduced by
 RT analysis of cDNA clones from human cell line KG-1.";
 RL DNA Res. 2:167-174(1995).
 RN [2]
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
 RX MEDLINE=98395067; PubMed=9726964;
 RA Bagrodia S., Taylor S.J., Jordan K.A., Van Aelst L., Cerione R.A.;
 RT "A novel regulator of p21-activated kinases.";
 RL J. Biol. Chem. 273:23633-23636(1998).
 CC -1- FUNCTION: Acts as a Rac1 guanine nucleotide exchange factor (GEF)
 CC and can induce membrane ruffling (By similarity).
 CC -1- SUBUNIT: Interacts with Pak kinases through the SH3 domain.
 CC -1- ALTERNATIVE PRODUCTS: At least 2 isoforms are produced by
 CC alternative splicing.
 CC -1- SIMILARITY: CONTAINS 1 DBL-HOMOLOGY (DH) DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D63476; BAA09763.1; -
 DR HSSP; P19174; 2HSP.
 DR Genew: HGNC:15607; ARHGEF7.
 DR MIM; 605477;
 DR InterPro; IPR001331; GDS_CDC24.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000219; RhGEF.
 DR InterPro; IPR001452; SH3.
 DR Pfam; PF00018; SH3; 1.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00621; RhGEF; 1.
 DR ProDom; PD000066; SH3; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00325; RhGEF; 1.
 DR SMART; SM00326; SH3; 1.
 DR PROSITE; PS00741; DH_1; FALSE-NEG.
 DR PROSITE; PS50010; DH_2; 1.
 DR PROSITE; PS50003; PH_DOMAIN; 1.
 DR PROSITE; PS50002; SH3; 1.
 KW Guanine-nucleotide releasing factor; SH3 domain; Alternative splicing.
 FT DOMAIN 6
 FT 93 65 SH3.
 FT 295 400 PH.
 FT DOMAIN 93
 FT 295 400 PH.
 SQ SEQUENCE 646 AA; 73140 MW; B5B5A83F0EBC28D2 CRC64;
 Query Match: 62.1%; Score 36; DB 1; Length 646;
 Best Local Similarity 60.0%; Pred. NO. 1.6e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 OY 1 QEWLKKLKK 10
 DB 388 QEWVHKLKK 397
 RESULT 37
 YIO4_YEAST STANDARD; PRT; 691 AA.
 ID YIO4_YEAST
 AC P40460;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Hypothetical 80.5 kDa protein in SLN1-RAD25 intergenic region.
 GN YII144W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RA Barrrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
 RA Churcher C.M., Connor R., Copsey T., Dunt S., Devlin K., Fraser A.,
 RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
 RA Louis E., Iye G., Moule S., Moule T., Odell C., Pearson D.,
 RA Rajandream M.A., Riles L., Riles N., Rowley N., Skelton J., Smith V.,
 RA Walsh S.V., Whitehead S.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: SOME, TO S.POMBE SPBC11C11.03.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----


```

DR EMBL: Z38059; CAA86134.1; -
DR PIR: S48390; S48390.
DR SGD: S0001406; T1D3.
DR Hypothetical protein.
SW SEQUENCE 691 AA; 80487 MW; 7BDC942227A8093 CRC64;
OY
Db 372 QEWPGKLEKM 381
OY 2 QEWLKKLKKM 11
||| ||:|
Db
Query Match 62.1%; Score 36; DB 1; Length 691;
Best Local Similarity 70.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0.

RESULT 38
AMPL_PLAFO STANDARD; PRT; 1056 AA.
ID AMPL_PLAFO
AC O96935;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE M1-family aminopeptidase (EC 3.4.11.-).
OS Plasmodium falciparum (isolate FCB1 of Columbia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxId:186763;
RN [1]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND DEVELOPMENTAL STAGE.
RX MEDLINE:99094503; PubMed:9879894;
RA Florent I., Dery 2., Allary M., Monsigny M., Mayer R., Schrevel J.;
RT "A Plasmodium falciparum aminopeptidase gene belonging to the M1
RT family of zinc-metallopeptidases is expressed in erythrocytic
RT stages."
RL M01. Biochem. Parasitol. 97:149-160(1998).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- DEVELOPMENTAL STAGE: Expressed in erythrocytic stages.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M1.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
-----
CC EMBL; Y09081; CAA70301.1; -
DR MEROPS; M01.0PM; -
DR InterPro: IPR001930; Ala.peptidase.
DR InterPro: IPR000130; Zn.Metpeptidase.
DR Pfam; PF01433; Peptidase_M1; 1.
DR PRINTS; PR00756; ALADIPASE.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR HydroLase; Metalloprotease; Aminopeptidase; Zinc;
KW Metal-binding.
FT METAL 467 467 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 468 468 BY SIMILARITY.
FT METAL 471 471 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 490 490 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 551 551 PROTON DONOR (POTENTIAL).
SQ SEQUENCE 1056 AA; 122669 MW; 994078262E288B58 CRC64;
Query Match 62.1%; Score 36; DB 1; Length 1056;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0.

OY 1 QEWLKKLKKM 10
||| ||:|
Db 931 QEWLKKTVSR 940

RESULT 39
3300_HUMAN

```

ID	P300_HUMAN	STANDARD:	PRF: 2414 AA.
AC	Q09472;		
DT	15-JUL-1998 (Rel. 36, Created)		
DR	15-JUL-1998 (Rel. 36, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	EIA-associated protein p300.		
CN	EP300 OR P300.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_taxid=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=95011587; PubMed=7523245;		
RA	Eckner R., Even M.E., Newsome D., Decaprio J.A.,		
RT	Lawrence J.B., Livingston D.M.;		
RT	"Molecular cloning and functional analysis of the adenovirus EIA-		
RT	associated 300-kD protein (p300) reveals a protein with properties of		
RL	a transcriptional adaptor."		
RL	Genes Dev. 8:869-884(1994).		
RN	[2]		
RP	INTERACTION WITH PCAF.		
RX	MEDLINE=96300317; PubMed=8684459;		
RA	Yang X.-J., Ogrzyko V.V., Nishikawa J.-I., Howard B.H., Nakatani Y.;		
RT	"A p300/CBP-associated factor that competes with the adenoviral		
RT	oncoprotein EIA.";		
NL	Nature 382:319-324(1996).		
CC	- FUNCTION: Probable transcriptional adapter required for the		
CC	activity of certain complex transcriptional regulatory elements.		
CC	May have a function in cell cycle regulation. Binds to and may be		
CC	involved in the transforming capacity of the adenovirus EIA		
CC	protein.		
CC	- SUBUNIT: Interacts with PCAF.		
CC	- SUBCELLULAR LOCATION: Nuclear.		
CC	- SIMILARITY: CONTAINS 1 BROMODOMAIN.		
CC	- SIMILARITY: CONTAINS 1 ZF-TYPE ZINC FINGER.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/)		
CC	or send an email to license@isb-sib.ch .		
CC	-----		
DR	EMBL: U01877; AAA18639.1; --		
DR	TRANSFAC: T01427; --		
DR	GeneW: HGNC:3373; EP300.		
DR	MIM: 602700; --		
DR	InterPro: IPR001487; Bromodomain.		
DR	InterPro: IPR003101; KIX.		
DR	InterPro: IPR000197; TAZ_finger.		
DR	InterPro: IPR000433; ZnF_ZZ.		
DR	pfam: PF00439; bromodomain; 1.		
DR	pfam: PF00569; ZZ; 1.		
DR	pfam: PF02135; ZF-TAZ; 2.		
DR	pfam: PF02172; KIX; 1.		
DR	PRINTS: PR00503; BROMODOMAIN.		
DR	SMART: SM00297; BROMO; 1.		
DR	SMART: SM00297; ZnF_ZZ; 1.		
DR	PROSITE: PS00633; BROMODOMAIN_1; 1.		
DR	PROSITE: PS00633; BROMODOMAIN_2; 1.		
DR	PROSITE: PS01357; ZF_ZZ_1; 1.		
DR	PROSITE: PS01357; ZF_ZZ_2; 1.		
KW	Transcription regulation; Nuclear protein; Bromodomain; Cell cycle;		
KW	Zinc finger.		
FT	DOMAIN 1..17		
FT	DOMAIN 1067..1139		
FT	DOMAIN 1572..1818		
FT	ZN_FING 1664..1707		
FT	DOMAIN 797..800		
FT	DOMAIN 1519..1526		
FT	DOMAIN 2066		
FT	POLY-GLU.		
FT	POLY-GLN.		
FT	DOMAIN		

```

FT  DOMAIN 2190 2195 POLY-GLN.
SQ  SEQUENCE 2414 AA; 264143 MW; 6BFF909EE4B9D693 CRC64;
Query Match
Best Local Similarity 62.1%; Score 36; DB 1; Length 2414;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKL 8
DB 1463 LOEWYK 1470

RESULT 40
CBP_MOUSE STANDARD; PRT; 2441 AA.
ID CBP_MOUSE
AC P45481;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE CREB-binding protein.
GN CREBBP OR CBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=94019866; PubMed=8413673;
RA Chivria J.C., Kwok R.P.S., Lamb N., Haglwaara M., Montminy M.R.,
RA Goodman R.H.;
RT "Phosphorylated CREB binds specifically to the nuclear protein CBP.";
RL Nature 365:855-859(1993).
CC -I- FUNCTION: MEDIATES CAMP-GENE REGULATION BY BINDING SPECIFICALLY TO
CC PHOSPHORYLATED CREB PROTEIN, ACTING AS A COACTIVATOR. CBP AUGMENTS
CC THE ACTIVITY OF PHOSPHORYLATED CREB TO ACTIVATE TRANSCRIPTION OF
CC CAMP-RESPONSIVE GENES.
CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- SIMILARITY: CONTAINS 1 ZINC-FINGER.
CC -I- SIMILARITY: CONTAINS 1 ZINC-FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: S66385; AAB28651.1; -
DR TRANSFAC: T01318; -
DR MGD: MGI:1098280; Crebbp.
DR InterPro: IPR001487; Bromodomain.
DR InterPro: IPR003101; KIX.
DR InterPro: IPR000197; TAF_finger.
DR InterPro: IPR000433; ZNF_ZZ.
DR Pfam: PF00439; bromodomain; 1.
DR Pfam: PF00569; Z2; 1.
DR Pfam: PF02135; zf-TAF; 2.
DR Pfam: PF02172; KIX; 1.
DR PRINTS: PR00503; BROMODOMAIN.
DR SMART: SM00291; BROMO; 1.
DR SMART: SM00291; ZNF_ZZ; 1.
DR PROSITE: PS00633; BROMODOMAIN_1; 1.
DR PROSITE: PS00614; BROMODOMAIN_2; 1.
DR PROSITE: PS01357; ZF_ZZ_1; 1.
DR PROSITE: PS01355; ZF_ZZ_2; 1.
DR Transcription regulation; Nuclear protein; Activator; Bromodomain;
KW Zinc-finger.
FT  DOMAIN 1104 1176 BROMODOMAIN.
FT  ZN_FING 1702 1745 -Z- TYPE.
FT  DOMAIN 1062 1065 POLY-GLU.
FT  DOMAIN 1556 1563 POLY-GLU.

```

```

FT  DOMAIN 1944 1949 POLY-PRO.
FT  DOMAIN 1968 1971 POLY-GLN.
FT  DOMAIN 2082 2086 POLY-GLN.
FT  DOMAIN 2200 2216 POLY-GLN.
FT  DOMAIN 2236 2239 POLY-GLN.
SQ  SEQUENCE 2441 AA; 265474 MW; 0AAB028C3112FA19 CRC64;
Query Match
Best Local Similarity 62.1%; Score 36; DB 1; Length 2441;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKL 8
DB 1500 LOEWYK 1507

RESULT 41
CBP_HUMAN STANDARD; PRT; 2442 AA.
ID CBP_HUMAN
AC Q92793; Q16376; O00147;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE CREB-binding protein.
GN CREBBP OR CBP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97385172; PubMed=9238046;
RA Schiblo O.M., Borrow J., Tomek R., Reshimi S., Harden A.,
RA Schlegelberger B., Housman D., Doggett N.A., Rowley J.D.,
RA Zelenik-Je N.J.;
RT "MLL is fused to CBP, a histone acetyltransferase, in therapy-related
RT acute myeloid leukemia with a t(11;16)(q23;p13.3).";
RL Proc. Natl. Acad. Sci. U.S.A. 94:8732-8737(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97321049; PubMed=9177780;
RA Giles R.H., Petrif F., Dauverse H.G., den Hollander A.I.,
RA Lushnikova T., van Ommeren G.J.B., Goodman R.H., Deaven L.L.,
RA Doggett N.A., Peters D.J.M., Breuning M.H.;
RT "Construction of a 1.2-kb contig surrounding, and molecular analysis
RT of, the human CREB-binding protein (CBP/CREBBP) gene on chromosome
RT 16p13.3.";
RL Genomics 42:96-144(1997).
RN [3]
RP SEQUENCE OF 1-405 FROM N.A.
RX MEDLINE=96376968; PubMed=8782817;
RA Borrow J., Stanton V.P., Andresen J.M., Beecher R., Behn F.G.,
RA Chaganti R.S.K., Civin C.I., Distche C., Dube I., Frischauf A.M.,
RA Housman D., Mittleman F., Volinia S., Watmore A.E., Housman D.E.;
RT "The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses
RT a putative acetyltransferase to the CREB-binding protein.";
RL Nat. Genet. 14:33-41(1996).
RN [4]
RP INTERACTION WITH PCAF.
RX MEDLINE=96300317; PubMed=8684459;
RA Yang X.-J., Ogryzko V.V., Nishikawa J.-I., Howard B.H., Nakatani Y.;
RT "A p300/CBP-associated factor that competes with the adenoviral
RT oncoprotein E1A.";
RL Nature 382:319-324(1996).
CC -I- FUNCTION: MEDIATES CAMP-GENE REGULATION BY BINDING SPECIFICALLY TO
CC PHOSPHORYLATED CREB PROTEIN, ACTING AS A COACTIVATOR. CBP AUGMENTS
CC THE ACTIVITY OF PHOSPHORYLATED CREB TO ACTIVATE TRANSCRIPTION OF
CC CAMP-RESPONSIVE GENES.
CC -I- SUBUNIT: Interacts with SMAD1, SMAD2, SMAD3 and PCAF.
CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- DISEASE: INVOLVED IN ACUTE LEUKEMIAS BY CHROMOSOMAL TRANSLOCATIONS
CC T(8;16)(p11;p13) INVOLVING CBP AND MOZ, AND T(11;16)(Q23;P13.3)
CC INVOLVING CBP AND MLL.

```

DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L24E.
GN RPL24E OR APE0070.
OS Aeropyrum pernix.
OC Archaea: Crenarchaeota: Thermoprotei: Desulfurococcales;
OC Desulfurococaceae: Aeropyrum.
OX NCBI_TaxID=36636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1.
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hirno Y., Horikawa H., Yamazaki S., Hatakeyama Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Akai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamaya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
"Complete genome sequence of an aerobic hyper-thermophilic
crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
RL
RL
CC - SIMILARITY: BELONGS TO THE L24E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its use
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>).
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: AP000064; BAA81393.1; -
DR InterPro: IPR000988; Ribosomal_L24E.
DR Pfam: PF01246; Ribosomal_L24e.1.
DR PROSITE, PS01073; RIBOSOMAL_L24E; FALSE_NEG.
KW RIBOSOMAL protein; Complete proteome.
SQ SEQUENCE 63 AA: 7329 MW: 400668806FADEA82 CRC64;

Query Match 60.3%; Score 35; DB 1; Length 63;
Best local similarity 62.5%; Pred. No. 22;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0.
OY 3 EWKKKK 10
Db 52 EWVKKVK 59

RESULT 43
ID CENA_HUMAN STANDARD; PRT; 140 AA.
AC P49450;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Centromere protein A (CENP-A) (Centromere autoantigen A).
GN CENPA.
OS Homo sapiens (Human).
OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
OC Mammalia: Eutheria: Primates: Catarrhini: Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95050934; PubMed=7962047;
RA Sullivan K.F., Hechenberger M., Masri K.;
RT "Human CENP-A contains a histone H3 related histone fold domain that
RT is required for targeting to the centromere.";
RL J. Cell Biol. 127:581-592(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

```

RN [3]
RP SEQUENCE OF 1-33 FROM N.A.
RA Shelby R.D., Vafa O., Sullivan K.F.;
RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: COULD ACT AS A CORE HISTONE NECESSARY FOR THE ASSEMBLY
CC OF CENTROMERES. MAY REPLACE ONE OR BOTH COPIES OF HISTONE H3 IN A
CC CERTAIN SET OF CENTROMERIC NUCLEOSOMES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DOMAIN: THE C-TERMINUS IS REQUIRED FOR CENTROMERIC LOCALIZATION.
CC -1- SIMILARITY: BELONGS TO THE HISTONE H3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U14518; AAA57416.1; -
DR EMBL: BC002703; AA02703.1; -
DR EMBL: U82609; AA847505.1; -
DR Genew: HGNC:1851; CENPA.
DR MIM: 117139; -
DR InterPro: IPR000164; Histone_H3.
DR InterPro: IPR004822; Histone_core.
DR Pfam: PF00125; histone_1.
DR SMART: PS00428; H3; 1.
DR PROSITE: PS00959; HISTONE_H3_2; 1.
DR Chromosomal protein; Nuclear protein; DNA-binding; Centromere;
KW Antigen.
FT DOMAIN
FT SEQUENCE 140 AA; 15990 MW; 11A28FEB54486489 CRC64;
SQ
Query Match
Best Local Similarity 60.3%; Score 35; DB 1; Length 140;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 2 QEWLKLKK 11
DB 45 QGWLKEIRKL 54
-----
RESULT 44
YEST YEAST
ID YE2 YEAST STANDARD: PRT: 231 AA.
AC P40007:
DR 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 01-FEB-1995 (Rel. 31, Last annotation update)
DE Hypothetical 26.9 kDa protein in MN1-PM140 intergenic region.
GN YER002W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berto A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U18778; AA864535.1; -
DR SGD: S000804; YER002W.
KW Hypothetical protein.
SQ SEQUENCE 231 AA; 26901 MW; B1D27F0B817ADED CRC64;
Query Match
Best Local Similarity 60.3%; Score 35; DB 1; Length 231;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 2 QEWLKLKK 10
DB 182 EEWLEKLKK 190
-----
RESULT 45
KGUA_MYCGE
ID KGUA_MYCGE STANDARD: PRT: 240 AA.
AC P47353:
DR 01-FEB-1996 (Rel. 33, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanylate kinase (EC 2.7.4.8) (GMP kinase).
GN GMP OR MG107.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
[1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small R.V., Sandusky M., Fuhmann J.L.,
RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
CC -1- FUNCTION: ESSENTIAL FOR RECYCLING GMP AND INDIRECTLY, CGMP.
CC -1- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE GUANYLATE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U39691; AAC71325.1; ALT_INIT.
DR HSSP: P15454; IGRY.
DR TIGR: MG107; -
DR InterPro: IPR00619; Guanylate_kin.
DR Pfam: PF00625; Guanylate_kin; 1.
DR PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
KW Transferrase; Kinase; ATP-binding; Complete proteome.
FT NP_BIND 63 Kinase; ATP (By SIMILARITY).
SQ SEQUENCE 240 AA; 27520 MW; 0664E11FE900935D CRC64;
Query Match
Best Local Similarity 70.0%; Score 35; DB 1; Length 240;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 QEWLKLKK 10
DB 180 QEWLKLKK 189

```

```

RESULT 46
NNMT_MOUSE STANDARD; PRT: 264 AA.
ID NNMT_MOUSE
AC 055239;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nicotinamide N-methyltransferase (EC 2.1.1.1).
GN NNMT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DA/2J, C57BL/6J, C57BR/CDJ, AKR/J, and C3H/HeJ; TISSUE=Liver;
RX MEDLINE=98124104; PubMed=9464457;
RA Yan L., Olteneas D.M., Craddock T.L., Weinshtilboun R.M.;
RT "Mouse liver nicotinamide N-methyltransferase: cDNA cloning,
RT expression, and nucleotide sequence polymorphisms.";
RL Biochem. Pharmacol. 54:1139-1149(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/HeJ;
RX MEDLINE=98392849; PubMed=9726248;
RA Yan L., Olteneas D.M., Kozak C.A., Weinshtilboun R.M.;
RT "Mouse nicotinamide N-methyltransferase gene: molecular cloning,
RT structural characterization, and chromosomal localization.";
RL DNA Cell Biol. 17:659-667(1998).
CC -1- FUNCTION: CATALYZES THE N-METHYLATION OF NICOTINAMIDE AND OTHER
CC PYRIDINES TO FORM PYRIDINIUM IONS. THIS ACTIVITY IS IMPORTANT
CC FOR BIOPANSEFORMATION OF MANY DRUGS AND XENOBIOTIC COMPOUNDS.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + nicotinamide -> S-
CC adenosyl-L-homocysteine + 1-methylnicotinamide.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE NNMT/PNMT/TEMT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U86105; AAB94331.1; -
DR EMBL: U86106; AAB94332.1; -
DR EMBL: U86108; AAB94334.1; -
DR EMBL: U86107; AAB94333.1; -
DR EMBL: AF044960; AAC77360.1; -
DR EMBL: AF029756; AAC77360.1; JOINED.
DR MGD: MGI:1099443; Nmt.
DR InterPro: IPR001601; Methyltransf.
DR InterPro: IPR000940; NNMT_PNMT_TEMTfm.
DR Pfam: PF01234; NNMT_PNMT_TEMT.1.
DR PROSITE: PS01100; NNMT_PNMT_TEMT.1.
DR KEGG: Tnsferase; Methyltransferase.
SQ SEQUENCE 264 AA; 29597 MW; DC3FE92738B7EC90 CRC64;

```

```

Query Match 60.3%; Score 35; DB 1; Length 264;
Best Local Similarity 85.7%; Pred. No. 92;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LGEWLKK 7
II:||||
Db 94 LQKWLKK 100

```

```

ID TEMT_MOUSE STANDARD; PRT: 264 AA.
AC P40936; O9CZ50;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thioether S-methyltransferase (EC 2.1.1.96) (TEMT).
GN TEMT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=95119077; PubMed=7819283;
RA Warner D.R., Mozier N.M., Pearson J.D., Hoffman J.L.;
RT "Cloning and base sequence analysis of a cDNA encoding mouse lung
RT thioether S-methyltransferase.";
RL Biochim. Biophys. Acta 1246:160-166(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuohi P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barah G.,
RA Blake J., Boffelli D., Bojunga N., Carlini P., de Bonaide M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guelincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wyrshw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlski S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- FUNCTION: CATALYZES TRANSFER OF THE METHYL GROUP FROM S-
CC ADENOSYLMETHIONINE TO X IN COMPOUNDS OF THE STRUCTURE R-X-R',
CC WHERE X MAY BE SULFUR, SELENIUM, OR TELLURIUM, AND R AND R' MAY BE
CC VARIOUS ORGANIC GROUPS.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + dimethyl sulfide -
CC -1- SUBUNIT: BELONGS TO THE NNMT/PNMT/TEMT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M88694; AAA62365.1; -
DR EMBL: AK013010; BAB828594.1; -
DR MGD: MGI:102963; Temt.
DR InterPro: IPR001601; Methyltransf.
DR InterPro: IPR000940; NNMT_PNMT_TEMTfm.
DR Pfam: PF01234; NNMT_PNMT_TEMT.1.
DR PROSITE: PS01100; NNMT_PNMT_TEMT.1.
DR KEGG: Tnsferase; Methyltransferase.
SQ SEQUENCE 264 AA; 29460 MW; 58AC5BA580A8B2EE CRC64;

```

```

Query Match 60.3%; Score 35; DB 1; Length 264;
Best Local Similarity 85.7%; Pred. No. 92;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 LOEWLKK 7
 11:1111
 Db 95 LOKWLKK 101

RESULT 48
 AROC BUCAP STANDARD; PRT: 353 AA.
 ID AROC BUCAP
 AC 09ZHE9;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chorismate synthase (EC 4.6.1.4) (5-enolpyruvylshikimate-3-phosphate
 DE phospholigase).
 GN AROC.
 OS Buchnera aphidicola (subsp. Schizaphis graminum).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=98794;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98440331; PubMed=9767718;
 RA Clark M.A., Baumann L., Baumann P.;
 RT "Buchnera aphidicola (Aphid endosymbiont) contains genes encoding
 RT enzymes of histidine biosynthesis."
 RL Curr. Microbiol. 37:356-358(1998).
 CC -1- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =
 CC chorismate + phosphate.
 CC -1- COFACTOR: REDUCED FLAVIN (BY SIMILARITY).
 CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
 CC seventh step.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CHORISMATE SYNTHASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL; AF067228; AAC97352.1; -
 DR InterPro; IPR000453; Chorismate_synth.
 DR Pfam; PF01264; Chorismate_synth; 1.
 DR Prodom; PD002941; Chorismate_synth; 1.
 DR TIGRFAMs; TIGR00033; aroc; 1
 DR PROSITE; PS00787; CHORISMATE_SYNTHASE_1; 1.
 DR PROSITE; PS00786; CHORISMATE_SYNTHASE_2; 1.
 DR PROSITE; PS00789; CHORISMATE_SYNTHASE_3; 1.
 DR Lyase; Aromatic amino acid biosynthesis.
 SQ SEQUENCE 353 AA; 38941 MW; 7C28C7DFC3C8929 CRC64;

Query Match 60.3%; Score 35; DB 1; Length 353;
 Best Local Similarity 70.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKK 10
 11:111111
 Db 189 LELWLKK 198

RESULT 49
 IPSF HUMAN STANDARD; PRT: 406 AA.
 ID IPSF HUMAN
 AC P05154; O9UG30; Q07616;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Plasma serine protease inhibitor precursor (PCI) (Protein C inhibitor)
 DE (Plasminogen activator inhibitor-3) (PAI3) (Acrosomal serine protease
 DE inhibitor).
 GN SERPINAS OR PCI OR PLANH3 OR PROCI.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87109153; PubMed=3027058;
 RA Suzuki K., Deyashiki Y., Nishio K., Kurachi K., Akira M.,
 RA Yamamoto S., Hashimoto S.;
 RT "Characterization of a cDNA for human protein C inhibitor. A new
 RT member of the plasma serine protease inhibitor superfamily."
 RL J. Biol. Chem. 262:611-616(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91048502; PubMed=2173165;
 RA Meijers J.C.M., Chung D.W.;
 RT "Evidence for a glycine residue at position 316 in human protein C
 RT inhibitor."
 RL Thromb. Res. 59:389-393(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91332018; PubMed=1714450;
 RA Meijers J.C.M., Chung D.W.;
 RT "Organization of the gene coding for human protein C inhibitor
 RT (plasminogen activator inhibitor-3). Assignment of the gene to
 RT chromosome 14."
 RL J. Biol. Chem. 266:15028-15034(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94198434; PubMed=8148499;
 RA Hayashi T., Suzuki K.;
 RT "Gene organization of human protein C inhibitor, a member of SERPIN
 RT family proteins encoded in five exons."
 RL Int. J. Hematol. 58:213-224(1993).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Radtke K.P., Greengard J.S., Fernandez J.A., Villoutreix B.O.,
 RA Griffin J.H.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A., AND VARIANTS G-44; V-55; S-64; V-94; E-105; P-115
 RP AND R-217.
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
 RA Nickerson D.A.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RA Strausberg R.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE OF 27-406 FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=93229000; PubMed=8471250;
 RA Moore A., Penfold L.M., Johnson J.L., Latchman D.S., Moore H.D.;
 RT "Human sperm-egg binding is inhibited by peptides corresponding to
 RT core region of an acrosomal serine protease inhibitor."
 RL Mol. Reprod. Dev. 34:280-291(1993).
 RN [9]
 RP SEQUENCE OF 28-406 FROM N.A.
 RC TISSUE=Testis;
 RA Poustka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [10]
 RP SEQUENCE OF 20-39.
 RX MEDLINE=90085156; PubMed=2556811;
 RA Laurell M., Stenflo J.;
 RT "Protein C inhibitor from human plasma: characterization of native
 RT and cleaved inhibitor and demonstration of inhibitor complexes with
 RT plasma kallikrein."
 RL Thromb. Haemost. 62:885-891(1989).
 RN [11]

RP 3D-STRUCTURE MODELING.
 RA MEDLINE-91046026; PubMed-2172389;
 RA Kuhn L.A., Griffin J.H., Fisher C.L., Greengard J.S., Bouma B.N.,
 RA Espana F., Tainer J.A.;
 RT "Elucidating the structural chemistry of glycosaminoglycan
 RT recognition by protein C inhibitor."
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8506-8510(1990).
 CC -!- FUNCTION: INHIBITS ACTIVATED PROTEIN C AS WELL AS PLASMINOGEN
 CC ACTIVATORS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: SYNTHESIZED IN LIVER; SECRETED IN PLASMA.
 CC -!- DISEASE: A DEFICIENCY OF PROTEIN C INHIBITOR IS THE CAUSE OF A
 CC COMBINED FACTOR V/VIII DEFICIENCY DISEASE.
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----
 DR EMBL: M68516; AAA02811.1; -;
 DR EMBL: J02639; AAA35688.1; -;
 DR EMBL: S69366; AAB30461.1; -;
 DR EMBL: S69364; AAB30461.1; JOINED.
 DR EMBL: S69574; AAB30461.1; JOINED.
 DR EMBL: S69365; AAB30461.1; JOINED.
 DR EMBL: AF361796; AAK27240.1; -;
 DR EMBL: U35464; AAB60386.1; -;
 DR EMBL: BC008915; AAH08915.1; -;
 DR EMBL: S58545; AAB26244.2; -;
 DR EMBL: AL080185; CAB45766.1; -;
 DR PIR: A39339; A39339.
 DR PDB: 1PAI; 15-OCT-94.
 DR PDB: 2PAI; 15-OCT-94.
 DR Genew; HGNC:8723; SERPINA5.
 DR MIM: 601841; -;
 DR InterPro; IPR000215; Serpin.
 DR Pfam; PF000079; serpin; 3.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Serpin: Serine protease inhibitor; Plasma; Glycoprotein; Signal;
 KW Polymorphism; 3D-structure.
 FT SIGNAL 1 19
 FT CHAIN 20 406 PLASMA SERINE PROTEASE INHIBITOR.
 FT CARBOHD 249 249 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT ACT_SITE 373 374 REACTIVE BOND.
 FT VARIANT 44 44 S -> G.
 FT VARIANT 55 55 /FTID-VAR_013080.
 FT VARIANT 64 64 A -> V (IN ALLELE).
 FT VARIANT 64 64 /FTID-VAR_007100.
 FT VARIANT 94 94 N -> S.
 FT VARIANT 94 94 /FTID-VAR_013081.
 FT VARIANT 105 105 G -> V.
 FT VARIANT 115 115 /FTID-VAR_013082.
 FT VARIANT 115 115 K -> E (IN ALLELE).
 FT VARIANT 115 115 /FTID-VAR_007101.
 FT VARIANT 115 115 L -> P.
 FT VARIANT 115 115 /FTID-VAR_013083.
 FT VARIANT 217 217 G -> R.
 FT VARIANT 217 217 /FTID-VAR_013084.
 FT CONFLICT 28 28 K -> E (IN REF. 9).
 FT CONFLICT 221 221 Q -> L (IN REF. 8).
 FT CONFLICT 335 335 G -> R (IN REF. 1 AND 8).
 FT CONFLICT 384 384 F -> S (IN REF. 8).
 SQ SEQUENCE 406 AA; 45701 MW; 0B9D1A519341B8C9 CRC64;

Query Match Score 35; DB 1; Length 406;
 Best Local Similarity 60.3%; Pred. No. 1.4e+02;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LOEWLKKLK 10
 Db 287 LKKWLKKRK 296
 RESULT 50
 YC46_ODOST
 ID YC46_ODOST STANDARD; PRT; 497 AA.
 AC P49541;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Hypothetical 56.9 kDa protein ycf46 (ORF497).
 GN YCF46.
 OS Odontella sinensis (Marine centric diatom).
 OG Chloroplast.
 OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
 OC Bidulphiophycidae; Eupodiscaceae; Odontella.
 OX NCBI_TaxID=2839;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kowalik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
 RT "The chloroplast genome of a chlorophyll a+c-containing alga,
 RT Odontella sinensis."
 RL Plant Mol. Biol. Rep. 13:336-342(1995).
 CC -!- SIMILARITY: DISTANTLY RELATED TO THE AAA FAMILY OF ATPASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----
 DR EMBL: Z67753; CAA91659.1; -;
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003959; AAA_ATPase_cent.
 DR Pfam; PF00004; AAA; 1.
 DR SMART; SM00382; AAA; 1.
 KW Chloroplast; Hypothetical protein; ATP-binding.
 FT NP_BIND 266 273 ATP (POTENTIAL).
 SQ SEQUENCE 497 AA; 56926 MW; EA08C08B95BA52CF CRC64;

Query Match Score 35; DB 1; Length 497;
 Best Local Similarity 66.7%; Pred. No. 1.7e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LOEWLKKLK 9
 Db 236 LKDWLKKRK 244
 Search completed: July 11, 2003, 10:49:02
 Job time : 33 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 10:46:43 ; Search time 77 Seconds
(without alignments)
29.435 Million cell updates/sec

Title: US-09-828-592-10
Perfect score: 58
Sequence: 1 LQEWLKKLKKM 11

Scoring table: BLASTSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 300 summaries

Database :

SEPREMBL_21.*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mmc:*
8: sp-organella:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-virus:*
16: sp-bacteriap:*
17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	42	72.4	118 16 067859	067859 aquifex aeo
2	42	72.4	375 17 087M1	087M1 methanosarc
3	42	72.4	434 17 087T18	087T18 methanosarc
4	42	72.4	434 17 087N0	087N0 methanosarc
5	42	72.4	434 17 087M5	087M5 methanosarc
6	42	72.4	434 17 087M0	087M0 methanosarc
7	42	72.4	434 17 087H32	087H32 methanosarc
8	41	70.7	263 2 09AP08	09AP08 uncultured
9	41	70.7	429 16 08X827	08X827 uncultured
10	41	70.7	804 10 08SAX2	08SAX2 escherichia
11	40	69.0	51 2 049028	049028 mycoplasma
12	40	69.0	85 12 08OKW3	08OKW3 mycoplasma
13	40	69.0	91 4 09UE54	09UE54 ectocarpus
14	40	69.0	156 5 09U0M2	09U0M2 plasmidium
15	40	69.0	315 16 08YJ09	08YJ09 bruceella me
16	40	69.0	448 17 08U3C3	08U3C3 pyrococcus

17	40	69.0	457 17 058339	058339 pyrococcus
18	40	69.0	465 4 013815	013815 homo sapien
19	40	69.0	745 10 09M105	09M105 arabidopsis
20	40	69.0	816 16 098P52	098P52 thizobium 1
21	40	69.0	1004 10 09LGM1	09LGM1 oryza sativ
22	40	69.0	1120 10 081509	081509 arabidopsis
23	40	69.0	1138 16 09PR90	09PR90 ureaplasma
24	39	67.2	185 4 09BOY5	09BOY5 homo sapien
25	39	67.2	213 10 0940Y4	0940Y4 oryza sativ
26	39	67.2	227 2 09S489	09S489 prochloroco
27	39	67.2	314 4 09NT60	09NT60 homo sapien
28	39	67.2	336 2 09F085	09F085 lactobacilli
29	39	67.2	371 10 09LNA0	09LNA0 arabidopsis
30	39	67.2	433 2 09KIG8	09KIG8 bacillus su
31	39	67.2	436 2 0939C9	0939C9 bacillus su
32	39	67.2	454 16 09K9M1	09K9M1 bacillus ha
33	39	67.2	480 17 09UY01	09UY01 pyrococcus
34	39	67.2	560 4 096N17	096N17 homo sapien
35	39	67.2	629 4 09UFY1	09UFY1 homo sapien
36	39	67.2	823 16 024895	024895 helicobacte
37	39	67.2	928 16 082252	082252 salmonella
38	39	67.2	955 16 09K8Y5	09K8Y5 bacillus ha
39	39	67.2	1114 10 09C8S4	09C8S4 arabidopsis
40	39	67.2	1233 10 09FXG4	09FXG4 arabidopsis
41	39	67.2	1640 9 09AZS0	09AZS0 bacterioph
42	39	67.2	1640 16 09CF57	09CF57 lactococcus
43	39	67.2	1713 9 094MA1	094MA1 lactococcus
44	38	65.5	120 4 09NMU7	09NMU7 homo sapien
45	38	65.5	131 17 028728	028728 archaeoglob
46	38	65.5	143 17 09V184	09V184 pyrococcus
47	38	65.5	196 11 09D068	09D068 mus musculi
48	38	65.5	252 17 058933	058933 methanococ
49	38	65.5	278 15 09WE07	09WE07 human immun
50	38	65.5	283 17 096YR6	096YR6 sulfobius
51	38	65.5	334 10 09C806	09C806 arabidopsis
52	38	65.5	355 5 09R909	09R909 pseudomonas
53	38	65.5	358 5 09R7J3	09R7J3 leishmania
54	38	65.5	402 9 080218	080218 methanobact
55	38	65.5	429 2 09AP59	09AP59 uncultured
56	38	65.5	464 4 09P2T2	09P2T2 homo sapien
57	38	65.5	473 4 09P2T3	09P2T3 homo sapien
58	38	65.5	521 16 08RAP5	08RAP5 thermotaneer
59	38	65.5	537 5 09U3Z8	09U3Z8 entamoeba h
60	38	65.5	561 10 09LITN0	09LITN0 arabidopsis
61	38	65.5	583 5 044468	044468 caenorhabdi
62	38	65.5	613 4 095942	095942 homo sapien
63	38	65.5	762 2 093584	093584 xanthomonas
64	38	65.5	786 16 032044	032044 bacillus su
65	38	65.5	859 15 090202	090202 human immun
66	38	65.5	1169 10 0947C8	0947C8 triticum mo
67	38	65.5	4131 5 019542	019542 caenorhabdi
68	37	63.8	117 16 0928R4	0928R4 listeria in
69	37	63.8	117 16 08Y4R5	08Y4R5 listeria mo
70	37	63.8	173 2 092AT6	092AT6 streptococ
71	37	63.8	218 16 09X121	09X121 thermotoga
72	37	63.8	233 16 08RDY5	08RDY5 fusobacteri
73	37	63.8	252 16 067070	067070 aquifex aeo
74	37	63.8	370 2 0919L7	0919L7 pasteurella
75	37	63.8	415 16 08YR34	08YR34 anabaena sp
76	37	63.8	429 2 09APP2	09APP2 uncultured
77	37	63.8	445 3 098776	098776 schizosacch
78	37	63.8	455 17 08U1K2	08U1K2 pyrococcus
79	37	63.8	550 16 08YR96	08YR96 anabaena sp
80	37	63.8	578 16 09P0K5	09P0K5 ureaplasma
81	37	63.8	703 3 09HEX5	09HEX5 pneumocysti
82	37	63.8	703 3 09C3Z3	09C3Z3 pneumocysti
83	37	63.8	724 16 09AD25	09AD25 streptomyce
84	37	63.8	731 16 08XUG9	08XUG9 raietonia s
85	37	63.8	751 5 002298	002298 caenorhabdi
86	37	63.8	816 2 09P5C5	09P5C5 agrobacteri
87	37	63.8	816 16 0989J2	0989J2 thizobium 1
88	37	63.8	818 2 093UW9	093UW9 agrobacteri
89	37	63.8	822 2 066172	066172 agrobacteri

90	37	63.8	822	16	09R488	09r488 agrobacteri	163	35	60.3	220	5	043991	043991 dictyosteli
91	37	63.8	852	2	P71180	P71180 enterobacte	164	35	60.3	231	6	028500	028500 macaca mula
92	37	63.8	852	2	052363	052363 escherichia	165	35	60.3	243	17	0803H6	0803H6 pyrococcus
93	37	63.8	852	2	09AHM4	09ahm4 comanonas t	166	35	60.3	273	15	09W0E6	09w0E6 human immun
94	37	63.8	852	2	08RSJ3	08rsj3 uncultured	167	35	60.3	276	15	09WDY6	09wdY6 human immun
95	37	63.8	942	10	09LGR9	09lgr9 oryza sativ	168	35	60.3	277	15	09WDY5	09wdY5 human immun
96	37	63.8	981	3	074419	074419 schizosacch	169	35	60.3	277	16	08ZK56	08zK56 salmonella
97	37	63.8	987	10	09LGL4	09lgl4 oryza sativ	170	35	60.3	278	15	09WEM9	09wem9 human immun
98	36	62.1	95	16	08R893	08r893 thermoanaer	171	35	60.3	278	15	09WEM0	09wem0 human immun
99	36	62.1	133	2	052285	052285 agrobacteri	172	35	60.3	278	15	09WEM1	09wem1 human immun
100	36	62.1	135	16	0930P9	0930P9 rhizobium m	173	35	60.3	278	15	09WEM4	09wem4 human immun
101	36	62.1	142	16	0922E3	0922E3 rhizobium m	174	35	60.3	279	15	09WEM6	09wem6 human immun
102	36	62.1	146	16	09CDI8	09cdi8 lactococcus	175	35	60.3	279	15	09WDM6	09wdm6 human immun
103	36	62.1	169	16	09HVM5	09hvm5 pseudomonas	176	35	60.3	280	15	09WED3	09wed3 human immun
104	36	62.1	175	4	096GX3	096gx3 homo sapien	177	35	60.3	290	10	08VXY4	08vxy4 human immun
105	36	62.1	175	17	08U021	08u021 pyrococcus	178	35	60.3	292	16	0980J1	098gJ1 mycoplasma
106	36	62.1	201	17	058670	058670 pyrococcus	179	35	60.3	299	16	0980J1	098gJ1 mycoplasma
107	36	62.1	229	5	08W053	08w053 geodia cydo	180	35	60.3	299	16	092564	092564 homo sapien
108	36	62.1	256	8	094Z15	094z15 schizophyll	181	35	60.3	313	10	09S9M7	09s9M7 arabisdopsis
109	36	62.1	275	16	08Y478	08y478 listeria mo	182	35	60.3	320	5	08WT03	08wt03 plasmodium
110	36	62.1	276	15	09WERS	09wers human immun	183	35	60.3	325	2	093GL1	093gl1 borrelia bu
111	36	62.1	279	11	09DA35	09da35 mus musculu	184	35	60.3	332	16	051011	051011 borrelia bu
112	36	62.1	287	16	0980X4	0980x4 mycoplasma	185	35	60.3	336	16	09WTF8	09wtf8 thermotoga
113	36	62.1	310	16	09CER1	09cer1 lactococcus	186	35	60.3	337	17	0980G7	0980g7 sulfolobus
114	36	62.1	318	4	09UPM2	09upm2 homo sapien	187	35	60.3	344	17	08TPH8	08tpH8 methanosarc
115	36	62.1	340	16	08XIB8	08xid8 clostridium	188	35	60.3	368	4	09H7I9	09h7I9 homo sapien
116	36	62.1	361	2	09RB22	09rbz2 pseudomonas	189	35	60.3	394	10	0940I8	0940I8 arabisdopsis
117	36	62.1	402	1	09HH69	09hh69 methanobact	190	35	60.3	398	17	09V0G5	09v0G5 pyrococcus
118	36	62.1	414	2	053680	053680 stigmatella	191	35	60.3	411	17	09UX68	09ux68 sulfolobus
119	36	62.1	424	16	08XG70	08xg70 salmonella	192	35	60.3	422	4	09GZK2	09gzK2 homo sapien
120	36	62.1	467	17	08TU02	08tu02 thermosarc	193	35	60.3	422	4	09GZK3	09gzK3 homo sapien
121	36	62.1	472	17	097C04	097c04 thermoplas	194	35	60.3	469	16	08RDZ6	08rdZ6 thermoanaer
122	36	62.1	479	16	097DB5	097db5 clostridium	195	35	60.3	472	12	09HIT2	09hit2 thermoplas
123	36	62.1	492	1	09HH59	09hh59 methanobact	196	35	60.3	474	12	055599	055599 garlic viru
124	36	62.1	498	16	099R30	099r30 streptococ	197	35	60.3	481	10	09MB76	09mb76 papulus eup
125	36	62.1	560	4	09U1Y0	09u1Y0 homo sapien	198	35	60.3	485	10	08R7I1	08r7I1 thermoanaer
126	36	62.1	611	3	094242	094242 schizosacch	199	35	60.3	499	16	09LR70	09lR70 arabisdopsis
127	36	62.1	726	16	08ZGS5	08zgs5 yersinia pe	200	35	60.3	545	3	059836	059836 schizosacch
128	36	62.1	729	5	09YRX3	09yrx3 caenorhabdi	201	35	60.3	556	16	053678	053678 mycobacteri
129	36	62.1	741	17	08TU06	08tus6 methanopyru	202	35	60.3	577	17	0976K9	0976K9 sulfolobus
130	36	62.1	800	2	08RPN2	08rpn2 ehrlichia c	203	35	60.3	590	16	P73561	P73561 synechocyst
131	36	62.1	801	16	08RPM6	08rpm6 anaplasma p	204	35	60.3	605	10	09SRZ8	09srZ8 arabisdopsis
132	36	62.1	805	16	092J06	092j06 rickettsia	205	35	60.3	611	10	09LKB4	09lKB4 arabisdopsis
133	36	62.1	805	16	09ZEA5	09zeA5 rickettsia	206	35	60.3	620	12	08QOC0	08qoc0 equus cabal
134	36	62.1	923	4	060424	060424 homo sapien	207	35	60.3	645	10	09FL05	09fl05 arabisdopsis
135	36	62.1	1085	10	09CA42	09ca42 arabisdopsis	208	35	60.3	703	5	09N5P3	09n5P3 caenorhabdi
136	36	62.1	1145	4	09UGR7	09ugr7 homo sapien	209	35	60.3	714	10	094HE6	094he6 oryza sativ
137	36	62.1	1161	4	09MYN5	09myn5 homo sapien	210	35	60.3	721	10	09M277	09m277 arabisdopsis
138	36	62.1	1170	4	09MYN6	09myn6 homo sapien	211	35	60.3	763	10	08S603	08s603 oryza sativ
139	36	62.1	1170	16	08YCA4	08yc44 bruceella me	212	35	60.3	793	10	08S774	08s774 oryza sativ
140	36	62.1	1198	4	09UEG3	09ueg3 homo sapien	213	35	60.3	816	16	08XW94	08xw94 ralstonia s
141	36	62.1	1245	3	003868	003868 saccharomyc	214	35	60.3	997	10	08SHP6	08shp6 oryza sativ
142	36	62.1	2404	5	09VE34	09ve34 drosophilla	215	35	60.3	1067	16	08XHB3	08xhb3 clostridium
143	36	62.1	2429	11	080ZV8	08qzv8 mus musculu	216	35	60.3	1121	16	09KOC0	09koc0 vibrio chol
144	36	62.1	2613	5	09GYD1	09gyd1 leishmania	217	35	60.3	1126	15	P90246	P90246 feline immu
145	36	62.1	3190	5	001368	001368 drosophilla	218	35	60.3	1266	16	08XMK8	08xmK8 clostridium
146	36	62.1	3275	10	09W321	09w321 drosophilla	219	35	60.3	1273	10	0941I9	0941I9 oryza sativ
147	35	60.3	69	10	09FPI5	09fpi5 arabisdopsis	220	35	60.3	1277	10	09AXD3	09axD3 zea mays (m
148	35	60.3	91	15	040282	040282 human immun	221	35	60.3	1278	10	08S455	08s455 zea mays (m
149	35	60.3	114	4	09BVM2	09bvm2 homo sapien	222	35	60.3	1278	10	08S453	08s453 zea mays (m
150	35	60.3	115	16	097MK1	097mk1 clostridium	223	35	60.3	1371	16	09JMS3	09jms3 escherichia
151	35	60.3	127	4	096FF4	096ff4 homo sapien	224	35	60.3	1543	12	080XJ1	080xJ1 garlic viru
152	35	60.3	129	11	08R565	08r565 cricetus	225	35	60.3	1544	12	080XJ1	080xJ1 garlic viru
153	35	60.3	134	16	09JYE4	09jye4 neisseria m	226	35	60.3	1550	12	055601	055601 garlic viru
154	35	60.3	134	16	09JTM8	09jtm8 neisseria m	227	35	60.3	1555	5	09U0N0	09u0N0 plasmodium
155	35	60.3	138	4	09H2A6	09h2a6 homo sapien	228	35	60.3	1557	12	08QXW5	08qxW5 garlic viru
156	35	60.3	159	2	050351	050351 lactobacilli	229	35	60.3	1625	12	055597	055597 garlic viru
157	35	60.3	166	5	097112	097112 dictyosteli	230	35	60.3	2514	5	09Y0G1	09y0G1 caenorhabdi
158	35	60.3	174	5	0966R2	0966r2 dictyosteli	231	35	60.3	2531	5	022258	022258 caenorhabdi
159	35	60.3	175	16	097DV1	097dv1 clostridium	232	34	58.6	76	12	098424	098424 paramacium
160	35	60.3	180	5	08USU5	08usu5 encephalito	233	34	58.6	78	11	08RIC8	08ric8 mus musculu
161	35	60.3	201	17	090U26	090uz6 pyrococcus	234	34	58.6	89	9	08SDV8	08sdv8 bacterioph
162	35	60.3	205	5	09VTA7	09vtat7 drosophilla	235	34	58.6	100	11	09CV35	09cv35 mus musculu

236	34	58.6	104	11	Q8R212	Q8R212 mus musculu
237	34	58.6	112	17	Q8RSS2	Q8RSS2 methanosarc
238	34	58.6	113	16	Q9Z8R2	Q9Z8R2 chlamydia p
239	34	58.6	131	16	Q67408	Q67408 aquifex aeo
240	34	58.6	142	5	Q01542	Q01542 caenorhabdi
241	34	58.6	151	17	Q97YU5	Q97YU5 sulfolobus
242	34	58.6	155	16	Q34694	Q34694 bacillus su
243	34	58.6	166	16	Q92FK8	Q92FK8 listeria in
244	34	58.6	170	10	Q9LYV6	Q9LYV6 arabidopsis
245	34	58.6	180	2	Q9EVX3	Q9EVX3 streptococc
246	34	58.6	190	3	Q07070	Q07070 saccharomyc
247	34	58.6	194	11	Q9UK74	Q9UK74 ratu mus nov
248	34	58.6	194	11	Q921E2	Q921E2 mus musculu
249	34	58.6	195	4	Q9HC00	Q9HC00 homo sapien
250	34	58.6	195	12	Q9IM24	Q9IM24 cactus viru
251	34	58.6	195	13	Q9IMW2	Q9IMW2 hydrangea r
252	34	58.6	197	13	Q90WX3	Q90WX3 brachydanio
253	34	58.6	197	13	Q8U0U0	Q8U0U0 brachydanio
254	34	58.6	197	13	Q8QFN0	Q8QFN0 fugu rubrip
255	34	58.6	201	11	Q8VBV8	Q8VBV8 mus musculu
256	34	58.6	203	16	Q9CG23	Q9CG23 lactococcus
257	34	58.6	207	12	Q8V9P9	Q8V9P9 sulfolobus
258	34	58.6	207	12	Q8QLA2	Q8QLA2 sulfolobus
259	34	58.6	209	16	Q9KMT0	Q9KMT0 vibrio chol
260	34	58.6	211	17	Q29130	Q29130 archaeoglob
261	34	58.6	221	11	Q9D1N4	Q9D1N4 mus musculu
262	34	58.6	223	5	Q9VN02	Q9VN02 dirosophila
263	34	58.6	223	10	Q9ZM26	Q9ZM26 arabidopsis
264	34	58.6	229	3	Q06070	Q06070 saccharomyc
265	34	58.6	229	4	Q8TD96	Q8TD96 homo sapien
266	34	58.6	233	5	Q9VWB1	Q9VWB1 dirosophila
267	34	58.6	246	16	Q97J53	Q97J53 clostridium
268	34	58.6	248	2	Q8RMB8	Q8RMB8 cytophaga j
269	34	58.6	248	11	Q9D423	Q9D423 mus musculu
270	34	58.6	250	5	Q17765	Q17765 caenorhabdi
271	34	58.6	252	12	Q91906	Q91906 influenza a
272	34	58.6	258	16	Q9KU02	Q9KU02 vibrio chol
273	34	58.6	259	17	Q9HM32	Q9HM32 thermoplasma
274	34	58.6	263	11	Q60407	Q60407 cricetus cr
275	34	58.6	264	2	Q59538	Q59538 mycoplasma
276	34	58.6	268	15	Q9WDC3	Q9WDC3 human immun
277	34	58.6	271	15	Q9WE02	Q9WE02 human immun
278	34	58.6	277	15	Q9WD27	Q9WD27 human immun
279	34	58.6	277	15	Q9WEP9	Q9WEP9 human immun
280	34	58.6	277	15	Q9WEQ4	Q9WEQ4 human immun
281	34	58.6	279	15	Q9WDC4	Q9WDC4 human immun
282	34	58.6	279	15	Q9WEP6	Q9WEP6 human immun
283	34	58.6	279	15	Q9WE01	Q9WE01 human immun
284	34	58.6	283	2	Q8S856	Q8S856 sphingomona
285	34	58.6	301	16	Q9X1D8	Q9X1D8 thermotoga
286	34	58.6	309	5	Q8SVT9	Q8SVT9 encephalito
287	34	58.6	316	1	Q93716	Q93716 methanosarc
288	34	58.6	316	17	Q8TU06	Q8TU06 methanosarc
289	34	58.6	323	16	Q9KSU3	Q9KSU3 vibrio chol
290	34	58.6	324	16	Q9PKC9	Q9PKC9 chlamydia m
291	34	58.6	324	16	Q84267	Q84267 chlamydia t
292	34	58.6	338	16	P73085	P73085 synecocyst
293	34	58.6	348	5	Q9Y906	Q9Y906 dirosophila
294	34	58.6	348	7	Q19481	Q19481 heterodontu
295	34	58.6	350	16	P96600	P96600 bacillus su
296	34	58.6	352	5	Q9VPE4	Q9VPE4 dirosophila
297	34	58.6	352	16	Q8YB70	Q8YB70 dirosophila
298	34	58.6	356	7	Q9XK31	Q9XK31 oryzias lat
299	34	58.6	362	2	Q9ZFP2	Q9ZFP2 bacillus me
300	34	58.6	371	16	Q8Y582	Q8Y582 listeria mo

ALIGNMENTS

RESULT 1
067859
ID 067859

PRELIMINARY: PRT: 118 AA.

AC 067859: 72.4% Score 42; DB 16; Length 118;
Best Local Similarity 70.0%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LEWIKLKK 10
1111111111
8 LSEWIKELKK 17

DB

Query Match
Best Local Similarity 70.0%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

AC 067859: 72.4% Score 42; DB 16; Length 118;
Best Local Similarity 70.0%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LEWIKLKK 10
1111111111
8 LSEWIKELKK 17

DB

Query Match
Best Local Similarity 70.0%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

RESULT 2
Q8TM41
ID Q8TM41
AC Q8TM41
AD Q8TM41
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Predicted protein.
GN MA2829.
OS Methanosarcina acetivorans.
OC Archaeae; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., MacDonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atwood D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Deatellano K., Johnson R.,
RA Lincon L., McEwan P., McKernan K., Talamas J., Tjirelli A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuestner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010983; AAM06208.1; -
KW Complete proteome.
SQ SEQUENCE 375 AA: 45263 MW: 870342B4C45F3782 CRC64;

Query Match 72.4% Score 42; DB 17; Length 375;
Best Local Similarity 77.8%; Pred. No. 68;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 OEMLKLRK 10
 |||:| |
 DB 294 OEMLKLRK 302

RESULT 3

OBTN8 PRELIMINARY: PRT: 434 AA.

AC OBTN8: 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DE Predicted protein.

OS Methanosarcina acetivorans.
 OC Archaea: Euryarchaeota; Methanococci; Methanosarcinales;
 OX NCBI_TaxID=2214;

RP SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=11932238;

RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., Deatellano K., Johnson R.,
 RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;

RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 and physiological diversity."

RL Genome Res. 12:532-542(2002).

DR EMBL: AE010699; AAM03840.1; -.

KW Complete proteome.

SO SEQUENCE 434 AA; 52162 MW; 7D097E0BC648DB5 CRC64;

Query Match 72.4%; Score 42; DB 17; Length 434;

Best Local Similarity 77.8%; Pred. No. 78;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 OEMLKLRK 10
 |||:| |
 DB 294 OEMLKLRK 302

RESULT 4

OBTN8 PRELIMINARY: PRT: 434 AA.

AC OBTN8: 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DE Predicted protein.

OS Methanosarcina acetivorans.
 OC Archaea: Euryarchaeota; Methanococci; Methanosarcinales;
 OX NCBI_TaxID=2214;

RP SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=11932238;

RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., Deatellano K., Johnson R.,
 RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,

RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;

RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 and physiological diversity."

RL Genome Res. 12:532-542(2002).

DR EMBL: AE010942; AAM05880.1; -.

KW Complete proteome.

SO SEQUENCE 434 AA; 52274 MW; 6A2501E74F8791E CRC64;

Query Match 72.4%; Score 42; DB 17; Length 434;

Best Local Similarity 77.8%; Pred. No. 78;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 OEMLKLRK 10
 |||:| |
 DB 294 OEMLKLRK 302

RESULT 5

OBTN8 PRELIMINARY: PRT: 434 AA.

AC OBTN8: 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DE Predicted protein.

OS Methanosarcina acetivorans.
 OC Archaea: Euryarchaeota; Methanococci; Methanosarcinales;
 OX NCBI_TaxID=2214;

RP SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=11932238;

RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., Deatellano K., Johnson R.,
 RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;

RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 and physiological diversity."

RL Genome Res. 12:532-542(2002).

DR EMBL: AE010951; AAM05948.1; -.

KW Complete proteome.

SO SEQUENCE 434 AA; 52232 MW; 084DA002A26E0DB7 CRC64;

Query Match 72.4%; Score 42; DB 17; Length 434;

Best Local Similarity 77.8%; Pred. No. 78;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 OEMLKLRK 10
 |||:| |
 DB 294 OEMLKLRK 302

RESULT 6

OBTN8 PRELIMINARY: PRT: 434 AA.

AC OBTN8: 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DE Predicted protein.

GN MA2626.
 OC Methanosarcina acetivorans.
 OC Archaea: Euryarchaeota: Methanococci: Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C2A / ATCC 35395 / DSM 2834;
 RA MEDLINE=21929760; PubMed=11932238;
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
 RA Linton L., McKernan K., Talamas J., Titrrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.:
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 RT and physiological diversity."
 RL Genome Res. 12:532-542(2002).
 DR EMBL: AE010957; AAM06004.1; -
 KW Complete proteome.
 SQ SEQUENCE 434 AA; 52273 MW; C08FEE1E74F87A7E CRC64;
 Query Match 72.4%; Score 42; DB 17; Length 434;
 Best Local Similarity 77.8%; Pred. No. 78;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 2 OEMLKKLK 10
 DB 294 OEMLKKKK 302
 RESULT 7
 OQTH32
 ID 08TH32 PRELIMINARY: PRT: 434 AA.
 AC 08TH32;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Predicted protein.
 GN MA0149 OR MA2087 OR MA2725.
 OS Methanosarcina acetivorans.
 OC Archaea: Euryarchaeota: Methanococci: Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C2A / ATCC 35395 / DSM 2834;
 RA MEDLINE=21929760; PubMed=11932238;
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
 RA Linton L., McKernan K., Talamas J., Titrrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.:
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 RT and physiological diversity."
 RL Genome Res. 12:532-542(2002).
 DR EMBL: AE010957; AAM05487.1; -
 DR EMBL: AE010971; AAM06104.1; -
 KW Complete proteome.
 SQ SEQUENCE 434 AA; 52261 MW; D0710E02B9E4AB3 CRC64;

Query Match 72.4%; Score 42; DB 17; Length 434;
 Best Local Similarity 77.8%; Pred. No. 78;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 2 OEMLKKLK 10
 DB 294 OEMLKKKK 302
 RESULT 8
 O9AP08
 ID 09AP08 PRELIMINARY: PRT: 263 AA.
 AC 09AP08;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE DAPA-aminotransferase (Fragment).
 GN B10A.
 OS uncultured bacterium pCOSH1.
 OC Bacteria; environmental samples.
 OX NCBI_TaxID=143796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20575196; PubMed=11133432;
 RA Entcheva P., Liebl W., Johann A., Hartsch T., Strelt W.R.:
 RT "Direct cloning from enrichment cultures, a reliable strategy for
 RT isolation of complete operons and genes from microbial consortia."
 RL Appl. Environ. Microbiol. 67:89-99(2001).
 DR EMBL: AF250769; AAG60562.1; -
 DR HSSP: P04181; 20AT.
 DR InterPro: IPR000954; AminoLtran_3.
 DR Pfam: PF00202; aminoLtran_3; 1.
 DR PROSITE: PS00600; AA_TRANSFER_CLASS_3; 1.
 KW AminoLtransferase; Transferase.
 FT NON_TER
 SQ SEQUENCE 263 AA; 28566 MW; 9D5B166EFED0B11C CRC64;
 Query Match 70.7%; Score 41; DB 2; Length 263;
 Best Local Similarity 66.7%; Pred. No. 71;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 OY 3 EWLKKLKK 11
 DB 60 EWLKKIRKK 68
 RESULT 9
 O8X827
 ID 08X827 PRELIMINARY: PRT: 429 AA.
 AC 08X827;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE 7,8-diaminopelargonic acid synthetase.
 GN B10A OR Z0993 OR ECS0852.
 OS Escherichia coli O157:H7.
 OC Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
 RA MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.:
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
 RL Nature 409:529-533(2001).
 RN [2]
 SQ SEQUENCE FROM N.A.

RC STRAIN-0157:H7 / RIMD 05099932;
RX MEDLINE+21156231; PubMed+11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
Ra Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
Ra Iida T., Takami H., Honda T., Sasaki A. C., Ogasawara N., Yasunaga T.,
Ra Kudara S., Shiba T., Hattori M., Shingara H.;
RT "Complete genome sequence of enterohemorrhagic *Escherichia coli*
RT 0157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
DR EMBL; AE005258; AAG53145.1; -;
DR EMBL; AP002553; BA834275.1; -;
DR InterPro; IPR000954; AminoTran_3.
DR Pfam; PF00202; aminoTran_3; 1.
DR TIGRFAMs; TIGR00508; bioa; 1.
DR PROSITE; PS00600; AA_TRANSFER_CLASS_3; 1.
KW Complete proteome.
SQ SEQUENCE 429 AA: 47282 MW: 1857B9A55E1DE9A6 CRC64;

Query Match	70.7%	Score 41	DB 16	Length 429
Best Local Similarity	66.7%	Pred No. 1.1e+02		
Matches 6	Conservative 3	Mismatches 0	Indels 0	Gaps 0

```
QY      3 EWLKLLKKM 11
        |||:::|
Db     226 EWLKRIRKM 234
```

RESULT 10

ID	Q8SAX2	PRELIMINARY;	PRT;	804	AA.
----	--------	--------------	------	-----	-----

DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Putative NBS-LRR type resistance protein.
 GN OSJNB0042K08.1
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
 RA Sasaki C., Henry D., Oates R., Simmons J.,
 RT "Rice Genomic Sequence."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 QO EMBL: AC097447; AAL71135.1; -
 QO SEQUENCE 804 AA; 91592 MW; 35EBE83FE3C6A17 CRC64;

Query Match	70.7%	Score 41	DB 10	Length 804
Best Local Similarity	77.8%	Pred. No. 2e+02	0	Gaps 0
Matches 7; Conservative		2; Mismatches		

QY	1	LQENLKKLK	9
		1:111:111	
Db	66	LKENLRLKL	74

RESULT 11

ID Q49028 PRELIMINARY; PRT; 51 AA.

DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Asp-tRNA synthetase (Fragment).
OS *Mycoplasma capricolum*.
CC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes
OC Entomoplasmatales; Entomoplasmataceae.
OX NCBI_TaxID=2095;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 27343(KID);
RX MEDLINE=36059641; PubMed=7476192;
RA Bork P., Ouzounis C., Casati G., Schneider R., Sander C., Dolan M.,
RA Gilbert W., Gillevet P.M.;
RT "Exploring the *Mycoplasma capricolum* genome: a minimal cell reveals
its physiology";
RL Mol. Microbiol. 16:955-967(1995).

Query Match	69.0%;	Score 40;	DB 2;	Length 51;
Best Local Similarity	54.58;	Pred. No. 22;		
Matches 6; Conservative		Mismatches 4;	Indels 1;	Gaps 0

QY	1	LQEWLKKLKKM	11
		:: :	
Db	21	LQGWVKKIRKL	31

RESULT 12

ID	PRELIMINARY;	85 AA.
Q8QKW3	PRT;	

DT 01-JUN-2002 (TReMBrel. 21, Created)
DT 01-JUN-2002 (TReMBrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBrel. 21, Last annotation update)
DE Esv-1-33 precursor.
OS Ectocarpus siliculosus virus.
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phaeovirus.
OX NCBI_TaxId=37665;
RN [1]
RP
RN
RC SEQUENCE FROM N.A.
RC STRAIN=ESV-1;
RA Delaroque N., Bothe G., Pohl T., Knippers R., Mueller D., Boland W.,
RT "The complete nucleotide sequence of the Ectocarpus siliculosus virus
RT genome.";
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF204351; AKL14459.1; -.
KW Signal.
FT SIGNAL 1 24
FT POTENTIAL.
FT 85 AA; 10029 MW; 1BC357BB9630318F CRC64;
FT SEQUENCE

Query Match	69.08;	Score 40;	DB 12;	Length 85;
Best Local Similarity	66.78;	Pred. No. 36;		
Matches 6; Conservative		Mismatches 3;	Indels 0;	Gaps 0;

QY	3	EWLKKLKKM	11
		: : : :	
Db	77	EWVKRLKKL	85

RESULT 1.3

ID Q9UE54 PRELIMINARY; PRT; 91 AA.

DT 01-MAY-2000 (TREMBLrel. 13. Created)
DT 01-MAY-2000 (TREMBLrel. 13. Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19. Last annotation update)
DE Gene fragment for antithrombin-III (aa 262-353).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_taxid=9606;
[1]
PP SEQUENCE FROM N.A.
RX MEDLINE=86056311; PubMed=2998880;
RA Jägd S.; Vibe-pedersen K.; Magnusson S.;
RT "Location of two of the introns in the antithrombin-III gene.";

RL FEBS Lett. 193:213-216(1985).
 CC -1 SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR EMBL: X03203; CAA26951.1; -.
 DR HSSP: P01008; IATH.
 DR InterPro: IPR00215; Serpin.
 DR Pfam: PF00079; serpin; 1.
 DR Serpin.
 SQ SEQUENCE 91 AA; 10493 MW; C1A61BCB25A0875E CRC64;

Query Match 69.0%; Score 40; DB 4; Length 91;
 Best Local Similarity 63.6%; Pred. NO. 38;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 QEWLKKLKKM 11
 |||||:::|
 DB 43 LQEWLDELEEM 53

RESULT 14

O9U0M2 PRELIMINARY; PRT; 156 AA.

AC O9U0M2:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Partial, hypothetical protein, MAL1P3.15 (Fragment).
 GN MAL1P3.15.
 OS Plasmodium falciparum (isolate 3D7).
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3D7;
 RA Bowman S., Churcher C., Harris B., Harris D., Lawson D., Quail M.,
 RA Barrell B.; (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AL031746; CAB63570.1; -.
 DR EMBL: AL031746; CAB63570.1; -.
 FT Hypothetical protein.
 KW NON_TER 156 156
 SQ SEQUENCE 156 AA; 19180 MW; F7D8B6B75F0AE427 CRC64;

Query Match 69.0%; Score 40; DB 5; Length 156;
 Best Local Similarity 77.8%; Pred. NO. 62;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 QEWLKKLKK 10
 ::|||
 DB 12 KEWYKLLKK 20

RESULT 15

O8YJD9 PRELIMINARY; PRT; 315 AA.

AC O8YJD9:
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE INTEGRASE.
 GN BME10147.
 OS Brucella melitensis.
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
 RC MEDLINE=20020109; PubMed=11756688;
 RA DelVecchio V.G., Kapatal V., Redkar R.J., Patra G., Mujeer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 RA Jajlonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Leiteson J.-J.,
 RA Haselkorn R., Kyriades N., Overbeek R.;
 "The genome sequence of the facultative intracellular pathogen

RT Brucella melitensis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL: AE009457; AAL51329.1; -.
 DR InterPro: IPR002104; Phage_Integrase.
 DR InterPro: IPR004107; Phage_Integrase-N.
 DR Pfam: PF00589; Phage_Integrase; 1.
 DR Pfam: PF02899; Phage_Integrase-N; 1.
 DR Complete proteome.
 SQ SEQUENCE 315 AA; 34066 MW; C3D6A70583DD0692D CRC64;

Query Match 69.0%; Score 40; DB 16; Length 315;
 Best Local Similarity 70.0%; Pred. NO. 1.2e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 QEWLKKLKKM 11
 :|||:|
 DB 20 EEWLKSLEKM 29

RESULT 16

O8U3C3 PRELIMINARY; PRT; 448 AA.

AC O8U3C3:
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein PF0547.
 GN PF0547.
 OS Pyrococcus furiosus.
 CC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE010177; AAL80671.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 448 AA; 50756 MW; FB154F1722F57C8F CRC64;

Query Match 69.0%; Score 40; DB 17; Length 448;
 Best Local Similarity 54.5%; Pred. NO. 1.7e+02;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 QEWLKKLKKM 11
 :|||:|
 DB 436 LEWVKKCEKL 446

RESULT 17

O58339 PRELIMINARY; PRT; 457 AA.

AC O58339:
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein PH0578.
 GN PH0578.
 OS Pyrococcus horikoshii.
 CC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RC MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Hatakeyama Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa M., Ohnuki Y.,
 RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kishida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,

RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
DR EMBL: AP000002; BAA29667.1; -;
KW Hypothetical protein: Complete proteome.
SQ SEQUENCE 457 AA; 51916 MW; 53CB683601E1EE CRC64;

Query Match 69.0%; Score 40; DB 17; Length 457;
Best Local Similarity 54.5%; Pred. No. 1.7e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Oy 1 LOEWLKLKLM 11
ID 013815 PRELIMINARY; PRT; 465 AA.
AC 013815; P78439; P78447;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Antithrombin III.
GN A73.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tsuji H., Tanaka O., Nakagawa M., Tanaka S., Hashimoto-Gotoh T.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 138-255 FROM N.A.
RX MEDLINE=89050967; PubMed=3191114;
RA Bock S.C., Marthan J.A., Radziejewska E.;
RT "Antithrombin III from: proline-407 to leucine mutation in a highly
RT conserved region near the inhibitor reactive site.";
RL Biochemistry 27:6171-6178(1988).
RN [3]
RP SEQUENCE OF 1-14 FROM N.A.
RX MEDLINE=84169500; PubMed=6672771;
RA Bock S.C., Levitan D.J.;
RT "Characterization of an unusual DNA length polymorphism 5' to the
RT human antithrombin III gene.";
RL Nucleic Acids Res. 11:8569-8582(1983).
RN [4]
RP SEQUENCE OF 1-13 FROM N.A.
RA Rosenberg J.B., Amrani D.L., Bergtrom G.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-13 FROM N.A.
RX MEDLINE=85080031; PubMed=6096369;
RA Prochownik E.V., Orkin S.H.;
RT "In vivo transcription of a human antithrombin III 'mlnigene'.";
RL J. Biol. Chem. 259:15386-15392(1984).
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL: M21644; AA51794.1; -;
DR EMBL: M21643; AA51794.1; JOINED.
DR EMBL: M21643; AA51793.1; -;
DR EMBL: D29832; BAA06212.1; -;
DR EMBL: X00238; CAA25060.1; -;
DR EMBL: X00237; CAA25059.1; -;
DR EMBL: U11270; AAI1930.1; -;
DR HSSP: P01008; ZANT.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin. 1.
DR SMART: SMO0093; SERPIN. 1.
KW Serpin; Signal.
SQ SEQUENCE 465 AA; 52691 MW; C69214FCA12766D0 CRC64;

Query Match 69.0%; Score 40; DB 4; Length 465;
Best Local Similarity 63.6%; Pred. No. 1.7e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 LOEWLKLKLM 11
ID 09M105 PRELIMINARY; PRT; 745 AA.
AC 09M105;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative transposon protein.
GN A74G03860.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL161498; CAB60809.1; -;
DR InterPro: IPR004312; ATHILA.
DR InterPro: IPR001064; ATHILA.
DR Pfam: PF03078; ATHILA. 1.
DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN.1.
SQ SEQUENCE 745 AA; 85773 MW; CAB6C2EFC879871 CRC64;

RESULT 19

Query Match 69.0%; Score 40; DB 10; Length 745;
Best Local Similarity 63.6%; Pred. No. 2.7e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 LOEWLKLKLM 11
ID 098P52 PRELIMINARY; PRT; 816 AA.
AC 098P52;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Conjugal transfer protein, Tribe.
GN M195608.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti.";

RESULT 20

Query Match 69.0%; Score 40; DB 10; Length 745;
Best Local Similarity 63.6%; Pred. No. 2.7e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 LOEWLKLKLM 11
ID 098P52 PRELIMINARY; PRT; 816 AA.
AC 098P52;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Conjugal transfer protein, Tribe.
GN M195608.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti.";

Query Match 69.0%; Score 40; DB 10; Length 745;
Best Local Similarity 63.6%; Pred. No. 2.7e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 LOEWLKLKLM 11
ID 098P52 PRELIMINARY; PRT; 816 AA.
AC 098P52;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Conjugal transfer protein, Tribe.
GN M195608.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti.";

Query Match 69.0%; Score 40; DB 10; Length 745;
Best Local Similarity 63.6%; Pred. No. 2.7e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 LOEWLKLKLM 11
ID 098P52 PRELIMINARY; PRT; 816 AA.
AC 098P52;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Conjugal transfer protein, Tribe.
GN M195608.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti.";

Db 1095 QEWLKVRS 1104

RESULT 24

ID Q9BOY5 PRELIMINARY; PRT; 185 AA.

AC Q9BOY5;

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DE D511824.2.5 (Phospholipase C, gamma 1 (Formerly subtype 148))

DE (1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase gamma 1,

DE PLC-148, EC 3.1.4.11), Isoform 5) (Fragment).

GN PLCG1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Bridgeman A.;

RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

DR EMBL; AL022394; CAC36283.1; -.

DR HSSP; P19174; 2HSP.

DR InterPro; IPR001849; PH.

DR InterPro; IPR001452; SH3.

DR Pfam; PF00169; PH; 1.

DR Pfam; PF00018; SH3; 1.

DR PRINTS; PR00452; SH3DOMAIN.

DR PRODOM; PD000066; SH3; 1.

DR SMART; SM00233; PH; 1.

DR SMART; SM00326; SH3; 1.

DR PROSITE; PS50003; PH DOMAIN; 1.

DR PROSITE; PS50002; SH3; 1.

KM SH3 domain.

FT NON_TER 1

SQ SEQUENCE 185 AA; 20718 MW; B3024E054FCF2D7 CRC64;

Query Match Best Local Similarity 67.2%; Score 39; DB 4; Length 185;

Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 QEWLKVRS 11

Db 130 QDMWKIRREV 140

RESULT 25

ID Q9ADY4 PRELIMINARY; PRT; 213 AA.

AC Q9ADY4;

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DE Putative vacuolar protein sorting-associated protein.

GN P0010B10.27.

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Eriactoidae; Oryzae; Oryza.

OX NCBI_TaxID=4530;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN-CV. NIPPONBARE;

RT Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC

clone:P0010B10.27";

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP003224; BAB63580.1; -.

SQ SEQUENCE 213 AA; 22983 MW; 4E78F90687C83828 CRC64;

Query Match 67.2%; Score 39; DB 10; Length 213;

Best Local Similarity 63.6%; Pred. No. 1.2e+02;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 QEWLKVRS 11

Db 169 VSEWLAKIKM 179

RESULT 26

ID Q9S489 PRELIMINARY; PRT; 227 AA.

AC Q9S489;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Hypothetical 27.2 kDa protein.

OS Prochlorococcus sp. (strain PCC 9511).

OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococaceae;

OC Prochlorococcus.

OX NCBI_TaxID=100363;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN-PCC 9511;

RX MEDLINE=21142517; PubMed=11208789;

RA Holtzendorff J., Partensky F., Jacquet S., Brynant F., Marie D.,

RA Gaczarek L., Mary I., Vaulot D., Hess W.R.;

RT "Diel Expression of Cell Cycle-Related Genes in Synchronized Cultures

of Prochlorococcus sp. Strain PCC 9511.";

RL J. Bacteriol. 183:915-920(2001).

DR EMBL; AF158628; AAD45693.1; -.

DR InterPro; IPR001173; Glycos_transf_2.

DR Pfam; PF00535; Glycos_transf_2; 1.

KM Hypothetical protein.

SQ SEQUENCE 227 AA; 27203 MW; 5E1E84F18559F019 CRC64;

Query Match Best Local Similarity 67.2%; Score 39; DB 2; Length 227;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 QEWLKVRS 11

Db 90 QDMWKIRREV 99

RESULT 27

ID Q9NT60 PRELIMINARY; PRT; 314 AA.

AC Q9NT60;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DE Hypothetical 36.1 kDa protein (Fragment).

GN DKFZP56401278.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA TISSUE-BRAIN;

RA Oltenswelder B., Obermaier B., Mewes H.W., Gassenhuber J., Wiemann S.;

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL137517; CAB70783.1; -.

KM Hypothetical protein.

FT NON_TER 1

SQ SEQUENCE 314 AA; 36150 MW; 7780E9C99A8A8F1B CRC64;

Query Match 67.2%; Score 39; DB 4; Length 314;

Best Local Similarity 60.0%; Pred. No. 1.7e+02;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QEWLKVRS 10

Db 11 QEWLKVRS 10

Db 1 LQOWIQRKSK 10

RESULT 28

ID 09F0B5 PRELIMINARY; PRT: 336 AA.

AC 09F0B5; .

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE Glycosyltransferase Epsm.

GN Epsm.

OS Lactobacillus delbrueckii (subsp. bulgaricus).

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;

OC Lactobacillaceae; Lactobacillus.

OX NCBI_TaxID=1585;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=LF15;

RA Lamothe G.T.;

RT "EPS cluster of Lactobacillus delbrueckii subsp. bulgaricus LF15.";

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF267127; AAC44717.1;

DR InterPro: IPR001173; Glycos_transf_2.

DR Pfam: PF00535; Glycos_transf_2; 1.

KW Transferase.

SO SEQUENCE 336 AA; 38480 MW; 83AE95B3B0DA7252 CRC64;

Query Match 67.2%; Score 39; DB 2; Length 336;

Best Local Similarity 87.5%; Pred. No. 1.8e+02;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WLKRLKRM 11

Db 323 WLKRLKSM 330

RESULT 29

ID 09LNAO PRELIMINARY; PRT: 371 AA.

AC 09LNAO; .

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE F5011.19.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C., Shinn P., Altafi H., Bel Q., Chin C., Chlou J., Choi E., Conn L., Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B., Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A., Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A., Ecker J.R.;

RA "Genomic sequence for Arabidopsis thaliana BAC F5011 from chromosome 1.";

RT RT

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Ecker J.R.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Ecker J.R.;

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,

RA Khan S., Kim C., Altafi H., Bel B., Chin C., Chlou J., Choi E., Conn L., Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B., Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A., Ecker J.;

RA Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC025416; AAF79638.1;

SO SEQUENCE 371 AA; 41266 MW; 6D0F8BC00565C17B CRC64;

Query Match 67.2%; Score 39; DB 10; Length 371;

Best Local Similarity 70.0%; Pred. No. 2e+02;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LQEWLKRK 10

Db 199 VQEWLKRYPK 208

RESULT 30

ID 09KIG8 PRELIMINARY; PRT: 433 AA.

AC 09KIG8; .

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Aspartate transaminase-like protein.

GN ORF.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacilliales;

OC Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=A13;

RA Conrad B., Feesche J., Steinborn G., Adler B., Hofemeister J.;

RT "The 1pa gene of Bacillus subtilis A13 affects a spectrum of antibiotic syntheses.";

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF233756; AAF87218.1;

DR InterPro: IPR004839; AminoTransf1/2.

DR Pfam: PF00155; aminotran_1_2; 1.

SO SEQUENCE 433 AA; 49103 MW; 08E698C0D8F2291F CRC64;

Query Match 67.2%; Score 39; DB 2; Length 433;

Best Local Similarity 75.0%; Pred. No. 2.3e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 QEWLKRK 9

Db 94 QEWLKRMR 101

RESULT 31

ID 0939C9 PRELIMINARY; PRT: 436 AA.

AC 0939C9; .

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Aspartate transaminase-like protein.

GN ASPB3.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacilliales;

OC Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=B3;

RA Yao S., Fuchsbaumer N., Hillen W.;

RT "Genetic loci required for biosynthesis of lipopeptide antibiotics Iturin A and surfactin in Bacillus subtilis B3.";

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY040867; AAL10665.1; -.
 DR InterPro: IPR004839; AminoTransfer1/2.
 DR Pfam: PF00155; aminotran_1.2; 1.
 SQ SEQUENCE 436 AA; 49117 MW; 42B3DD6B93B863E3 CRC64;

Query Match
 Best Local Similarity 67.2%; Score 39; DB 2; Length 436;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 EWLKRLKM 9
 |||||:
 DB 95 QEWLKKMR 102

RESULT 32

O9K9M1 PRELIMINARY: PRT: 454 AA.
 AC O9K9M1:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE AminoTransferase.
 GN BH2624.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis."
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL: AP001516; BAB06343.1; -.
 DR InterPro: IPR000954; AminoTran_3.
 DR Pfam: PF00202; aminotran_3; 1.
 DR TIGFAMS: TIGR00709; dat: 1.
 DR PROSITE: PS00600; AA_TRANSFER_CLASS_3; 1.
 KW Transferrase; Complete proteome.
 SQ SEQUENCE 454 AA; 49747 MW; 8C6590502AF8AF8E CRC64;

Query Match
 Best Local Similarity 67.2%; Score 39; DB 16; Length 454;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 EWLKRLKM 11
 |||||:
 DB 238 EWLKELRRM 246

RESULT 33

O9UYU1 PRELIMINARY: PRT: 480 AA.
 AC O9UYU1:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE Hypothetical protein PAB0938.
 GN PAB0938.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ORSAY;
 RA Hellig R.;
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome

RT structure and evolution.";
 RL Submitted (Jul-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AJ248287; CAB50321.1; -.
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 480 AA; 54696 MW; A16A07BDBDB97059 CRC64;

Query Match
 Best Local Similarity 67.2%; Score 39; DB 17; Length 480;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKKLKM 11
 |||||:
 DB 468 LEEWYKCKKL 478

RESULT 34

O96N17 PRELIMINARY: PRT: 560 AA.
 AC O96N17:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE CDNA FLJ31560 fis, clone NT2R12001409, weakly similar to
 DE dystrophin.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Niinomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
 RA Furuya T., Takahashi M., Kikawa E., Omura Y., Abe K., Kamihara K.,
 RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
 RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
 RA Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K., Isogai T.;
 RT "NEBO human cDNA sequencing project."
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AK056122; BAB71097.1; -.
 DR InterPro: IPR002017; Spectrin.
 DR Pfam: PF00435; Spectrin; 2.
 SQ SEQUENCE 560 AA; 64530 MW; 61FFCF86C191D4EB CRC64;

Query Match
 Best Local Similarity 67.2%; Score 39; DB 4; Length 560;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKKLKM 11
 |||||:
 DB 267 IQDWLSTKEM 277

RESULT 35

O9UFY1 PRELIMINARY: PRT: 629 AA.
 AC O9UFY1:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 71.4 kDa protein (Fragment).
 GN DKF2P434N101.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-TESTIS;
 RA Poustka A., Klein M., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.
 CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

DR EMBL: AL110247; CAB53696.1; -
DR HSSP: P08487; 2PLD.
DR InterPro: IPR000008; C2.
DR InterPro: IPR001849; PH.
DR InterPro: IPR001192; PL_PLC.
DR InterPro: IPR001711; PL_PLC_Y.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00168; C2; 1.
DR Pfam: PF00169; PH; 1.
DR Pfam: PF00387; PI-PLC-Y; 1.
DR Pfam: PF00018; SH3; 1.
DR PRINTS: PR00390; PPHLPASEC.
DR PRINTS: PR00452; SH3DOMAIN.
DR ProDom: PD000066; SH3; 1.
DR ProDom: PD000093; SH2; 1.
DR ProDom: PD001202; PL_PLC_Y; 1.
DR SMART: SM00239; C2; 1.
DR SMART: SM00233; PH; 1.
DR SMART: SM00149; PLCYC; 1.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PSS0004; C2_DOMAIN_2; 1.
DR PROSITE: PSS0003; PH_DOMAIN; 1.
DR PROSITE: PSS0008; PIPIC_Y_DOMAIN; 1.
DR PROSITE: PSS0002; SH3; 1.
DR Hypothetical protein: SH3 domain.
FT NON_TER 1
SQ SEQUENCE 629 AA; 71369 MW; 0DC47D6985BD95F CRC64;

Query Match 67.28; Score 39; DB 4; Length 629;
Best Local Similarity 45.5%; Pred. No. 3.3e+02;
Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOEWLKKLKK 11
DB 204 LQDWVKIREV 214

RESULT 36
ID 024895 PRELIMINARY; PRT; 823 AA.
AC 024895;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Modification methylase (EC 2.1.1.73) (Cytosine-specific
DE methyltransferase).
GN HP0034.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_Taxid=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2695 / ATCC 700392;
RX MEDLINE=97394467; Pubmed=9252185;
RA Tomb J.-F., White O., Keriavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Meldman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA CYTOSINE - S-
CC ADENOSYL-L-HOMOCYSTEINE + DNA 5-METHYLCYTOSINE.
CC -1- SIMILARITY: TO C-5 CYTOSINE-SPECIFIC DNA METHYLASE FAMILY.
DR EMBL: AE000527; AAD07124.1; -
DR HSSP: P14385; ZADM.

DR TIGR: HP0054; -
DR InterPro: IPR001525; C5_DNA_meth.
DR InterPro: IPR002052; N6_Mtase.
DR Pfam: PF00145; DNA_methylase; 1.
DR PROSITE: PS00094; C5_MTASE_1; UNKNOWN_1.
DR PROSITE: PS00092; N6_MTASE; UNKNOWN_1.
DR Hypothetical protein: Methyltransferase; Restriction system;
KW Transferase; Complete proteome.
SQ SEQUENCE 823 AA; 95114 MW; DEDAB2FEE67A6FD1 CRC64;

Query Match 67.28; Score 39; DB 16; Length 823;
Best Local Similarity 60.0%; Pred. NO. 4.2e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKKLKK 10
DB 246 MOEWIKOLKE 255

RESULT 37
ID 082252 PRELIMINARY; PRT; 928 AA.
AC 082252;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE DNA polymerase I.
GN SRY3881.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_Taxid=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; Pubmed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaiha M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltham T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagers K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
DR EMBL: AL627280; CAD03100.1; -
DR InterPro: IPR002562; 3_5_exonuclease.
DR InterPro: IPR002421; 5_3_exonuclease.
DR InterPro: IPR001098; DNA_POL.
DR InterPro: IPR002298; DNA_POL.
DR InterPro: IPR000513; EXO_N_1.
DR Pfam: PF01612; 3_5_exonuclease; 1.
DR Pfam: PF01617; 5_3_exonuclease; 1.
DR Pfam: PF02739; 5_3_exonuc_N; 1.
DR Pfam: PF00476; DNA_POL_A; 1.
DR PRINTS: PR00868; DNAPOL1.
DR SMART: SM00474; 35EXOC; 1.
DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00279; Hnh2; 1.
DR SMART: SM00482; POLA2; 1.
DR TIGRFAMS: TIGR00593; pola; 1.
DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
KW Complete proteome.
SQ SEQUENCE 928 AA; 103126 MW; 690BF98135C51262 CRC64;

Query Match 67.28; Score 39; DB 16; Length 928;
Best Local Similarity 60.0%; Pred. NO. 4.7e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKKLKK 10

Db 339 LESWIERLKK 348

RESULT 38

09K8M5 PRELIMINARY; PRT; 955 AA.
 AC 09K8M5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein BH2981.
 GN BH2981.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RA MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Ogasawara N., Kuhara S.,
 RA Fuji F., Hirama C., Nakamura Y.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL: AP001517; BAB06700.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 955 AA; 111131 MW; BCG9402HC822FE70 CRC64;

Query Match 67.2%; Score 39; DB 16; Length 955;
 Best Local Similarity 66.7%; Pred. No. 4.8e+02;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 QEWLKRKK 10
 :||||:|
 Db 748 REWLKRKK 756

RESULT 39

09C8S4 PRELIMINARY; PRT; 1114 AA.
 AC 09C8S4;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE DNA-directed RNA polymerase subunit, putative.
 GN T1P2.15.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gall J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luos J.S., Maiti R., Marzilli A.,
 RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salerg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utechtack T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,

RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:816-820(2000).
 DR EMBL: AC022455; AAG52049.1; -;
 DR InterPro: IPR001572; RNA_POL_B.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00562; RNA_POL_B; 2.
 DR PROSITE: PS01166; RNA_POL_BETA; 1.
 DR PROSITE: PS00030; RNA_RNP_1; UNKNOWN_1.
 SQ SEQUENCE 1114 AA; 125138 MW; 7B777562FE93F18 CRC64;

Query Match 67.2%; Score 39; DB 10; Length 1114;
 Best Local Similarity 54.5%; Pred. No. 5.5e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LOEWLKRKK 11
 :||||:|
 Db 328 LEWLKRCKSL 338

RESULT 40

09FXG4 PRELIMINARY; PRT; 1233 AA.
 AC 09FXG4;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE F1M18.2 protein.
 GN F1M18.2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
 RA Altafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
 RA Buehler E., Chao Q., Chin C., Chou J., Choi E., Gonzalez A.,
 RA Hwang B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
 RA Lenz C., Liu A., Liu S., Mukharzky N., Pham P., Sakano H., Shinn P.,
 RA Toriumi M., Vaysberg M., Yu G., Ecker J., Theologis A., Davis R.W.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC008030; AAG10602.1; -;
 DR InterPro: IPR001572; RNA_POL_B.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00562; RNA_POL_B; 3.
 DR PROSITE: PS01166; RNA_POL_BETA; 1.
 DR PROSITE: PS00030; RNA_RNP_1; UNKNOWN_1.
 SQ SEQUENCE 1233 AA; 138652 MW; 540549FEAEFA2A529 CRC64;

Query Match 67.2%; Score 39; DB 10; Length 1233;
 Best Local Similarity 54.5%; Pred. No. 6.1e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LOEWLKRKK 11
 :||||:|
 Db 418 LEWLKRCKSL 428

RESULT 41

09AZS0 PRELIMINARY; PRT; 1640 AA.
 AC 09AZS0;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Tail protein.
 GN ORF52.
 OS Bacteriophage b1L286.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
 OX NCBI_TaxID=151536;

```
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21113149; PubMed=11160885;
RA Chopin A., Bolotin A., Sorokin A., Ehrlich S.D., Chopin M.-C.;
RT "Analysis of six prophages in Lactococcus lactis IL1403: different
RT genetic structure of temperate and virulent phage populations.";
RL Nucleic Acids Res. 29:644-651(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Chopin A., Bolotin A., Sorokin A., Ehrlich S.D., Chopin M.-C.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF333669; AAK08339.1; -
SQ SEQUENCE 1640 AA; 176321 MW; CEE59A7A47B8F18 CRC64;

Query Match
Best Local Similarity 100.0%; Score 39; DB 9; Length 1640;
Pred. No. 7.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WLKRLK 10
DB 671 WLKRLK 677

RESULT 42
Q9CF57 PRELIMINARY; PRT; 1640 AA.
ID Q9CF57;
AC Q9CF57;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Prophage p13 protein 14.
GN P1314 OR L11388.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Manger S., Tallon O., Malarme K.,
RA Weissbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
DR EMBL; AE006370; AAK05486.1; -
KW Complete proteome.
SQ SEQUENCE 1640 AA; 176320 MW; CEE59A7A47B8F18 CRC64;

Query Match
Best Local Similarity 100.0%; Score 39; DB 16; Length 1640;
Pred. No. 7.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WLKRLK 10
DB 671 WLKRLK 677

RESULT 43
Q94MA1 PRELIMINARY; PRT; 1713 AA.
ID Q94MA1;
AC Q94MA1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 183.5 kDa protein.
OS Lactococcus phage BK5-T.
Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=31754;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21235794; PubMed=11336549;
RA Desiere F., Mahanivong C., Hillier A.J., Chandry P.S., Davidson B.E.,
```

```
RA Brussow H.;
RT "Comparative Genomics of Lactococcal Phages: Insight from the Complete
RT Genome Sequence of Lactococcus lactis Phage BK5-T.";
RL Virology 283:240-252(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Mahanivong C., Boyce J.D., Davidson B.E., Hillier A.J.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF176025; AAK56815.1; -
DR InterPro; IPR00189; SLT_domain.
DR Pfam; PF01464; SLT; 1.
KW Hypothetical protein.
SQ SEQUENCE 1713 AA; 183493 MW; DADDEB64DF875278 CRC64;

Query Match
Best Local Similarity 100.0%; Score 39; DB 9; Length 1713;
Pred. No. 8.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WLKRLK 10
DB 740 WLKRLK 746

RESULT 44
Q9NUM7 PRELIMINARY; PRT; 120 AA.
ID Q9NUM7;
AC Q9NUM7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE CDNA FLJ11266 fis, clone PLACE1009166.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Yamamoto S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Watanabe J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,
RA Niimiya K., Iwayanagi T.;
RT "NEO human CDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK002128; BAA92095.1; -
SQ SEQUENCE 120 AA; 14303 MW; C9605BF684726B2C CRC64;

Query Match
Best Local Similarity 63.6%; Score 38; DB 4; Length 120;
Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LOEWLKLK 11
DB 26 LOEWLKLK 36

RESULT 45
Q28728 PRELIMINARY; PRT; 131 AA.
ID Q28728;
AC Q28728;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Transcriptional regulatory protein, ARSR family.
GN AF1544.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
```

```

RP SEQUENCE FROM N.A.
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=96049343; PubMed=9389475;
RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervatage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodok A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., DAndrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
DR EMBL: AE000996; AAB89703.1; -
DR TIGR: AF1544; -
DR InterPro: IPR001845; HTH_ArsR.
DR SMART: SM00418; HTH_ARSR; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 131 AA; 15829 MW; EBB4CE972C4776A3 CRC64;

Query Match 65.5%; Score 38; DB 17; Length 131;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 2 QEWLKKKK 10
Db 32 KEWLKKLR 40

RESULT 46
Oy 09V184 PRELIMINARY; PRT; 143 AA.
AC 09V184;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein PAB2001.
GN PAB2001.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ORSAY;
RA Heilig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
RT structure and evolution."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ248284; CAB49466.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 143 AA; 15891 MW; B6142C8A37D3D399 CRC64;

Query Match 65.5%; Score 38; DB 17; Length 143;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LOEWLKKKK 10
Db 113 KEWLKKVEE 122

RESULT 47
Oy 09D068 PRELIMINARY; PRT; 196 AA.
AC 09D068;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

```

```

DE 2810027019Rik protein.
GN 2810027019Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL: AK011768; BAB27829.1; -
DR MGI: 1915451; 2810027019Rik.
SQ SEQUENCE 196 AA; 23071 MW; BF9124E270B590C8 CRC64;

Query Match 65.5%; Score 38; DB 11; Length 196;
Best Local Similarity 87.5%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 QEWLKKKK 9
Db 60 QEWLKKKK 67

RESULT 48
Oy 058933 PRELIMINARY; PRT; 252 AA.
AC 058933;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein MJ1538.
GN MJ1538.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanococcaceae; Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodok A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Kleink H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: TO YEAST KTI12 PROTEIN.
DR EMBL: U67594; AAB99557.1; -

```


RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Aikai A., Kosugi H., Hosogama A., Fukui S.

Search completed: July 11, 2003, 10:51:16
Job time : 90 secs

Search completed: July 11, 2003, 10:51:16
Job time : 90 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: July 11, 2003, 10:47:13 ; Search time 26 Seconds
(without alignments)
12.448 Million cell updates/sec

Title: US-09-828-592-10
Perfect score: 58
Sequence: 1 LQEWLKKKKM 11

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 300 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2-6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2-6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2-6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2-6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2-6/ptodata/1/1aa/PCTUS.COMB.pep:*
- 6: /cgn2-6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	69.0	464	1	US-08-046-431A-2
2	39	67.2	1290	1	US-08-138-641-2
3	39	67.2	1290	1	US-08-138-133-2
4	38	65.5	806	1	US-08-270-076A-11
5	37	63.8	429	3	US-08-411-768B-4
6	37	63.8	431	2	US-08-401-068-4
7	37	63.8	431	2	US-08-846-338-4
8	36	62.1	314	4	US-09-475-316A-118
9	36	62.1	560	4	US-09-134-001C-3153
10	36	62.1	2414	4	US-08-227-536-2
11	36	62.1	2414	5	PCT-US95-04682-2
12	36	62.1	2441	1	US-08-194-468-2
13	36	62.1	2441	3	US-08-961-739-2
14	36	62.1	2441	4	US-09-514-247A-8
15	36	62.1	2442	4	US-09-514-247A-10
16	35	60.3	16	1	US-07-789-179-2
17	35	60.3	16	1	US-08-240-712-9
18	35	60.3	16	1	US-08-443-890-9
19	35	60.3	16	1	US-08-444-942-2
20	35	60.3	16	1	US-08-446-105-2
21	35	60.3	16	1	US-08-444-939-2
22	35	60.3	16	2	US-08-444-991-2
23	35	60.3	16	2	US-08-450-733-2
24	35	60.3	16	4	US-09-058-562-9
25	35	60.3	16	4	US-08-444-915A-2
26	35	60.3	16	5	PCT-US92-09752-9
27	35	60.3	16	6	5304631-1

28	35	60.3	18	3	US-08-940-095-215	Sequence 215, App
29	35	60.3	18	3	US-08-940-093-215	Sequence 215, App
30	35	60.3	18	3	US-08-940-096-215	Sequence 215, App
31	35	60.3	18	4	US-09-465-719-215	Sequence 215, App
32	35	60.3	18	4	US-09-453-605-215	Sequence 215, App
33	35	60.3	18	4	US-09-453-838-215	Sequence 215, App
34	35	60.3	20	1	US-08-437-841-14	Sequence 14, App1
35	35	60.3	20	1	US-08-286-521-14	Sequence 14, App1
36	35	60.3	20	1	US-08-436-175-14	Sequence 14, App1
37	35	60.3	20	2	US-08-435-149-10	Sequence 10, App1
38	35	60.3	20	4	US-08-943-682-14	Sequence 14, App1
39	35	60.3	20	5	PCT-US95-09464-14	Sequence 14, App1
40	35	60.3	1024	4	US-09-562-737-85	Sequence 85, App1
41	35	60.3	1150	4	US-09-238-303-9	Sequence 9, App1
42	34	58.6	11	1	US-08-193-521-17	Sequence 17, App1
43	34	58.6	11	1	US-08-434-120-111	Sequence 111, App
44	34	58.6	11	1	US-08-465-325-110	Sequence 110, App
45	34	58.6	11	1	US-08-343-882-11	Sequence 11, App1
46	34	58.6	11	4	US-09-115-737-110	Sequence 110, App
47	34	58.6	12	1	US-08-465-325-133	Sequence 133, App
48	34	58.6	12	4	US-09-115-737-133	Sequence 133, App
49	34	58.6	14	1	US-08-193-521-19	Sequence 19, App1
50	34	58.6	14	1	US-08-434-120-113	Sequence 113, App
51	34	58.6	14	1	US-08-465-325-117	Sequence 117, App
52	34	58.6	14	4	US-09-115-737-117	Sequence 117, App
53	34	58.6	15	1	US-07-725-331-66	Sequence 66, App1
54	34	58.6	15	1	US-08-193-521-7	Sequence 7, App1
55	34	58.6	15	1	US-08-193-521-8	Sequence 8, App1
56	34	58.6	15	1	US-08-434-120-101	Sequence 101, App
57	34	58.6	15	1	US-08-434-120-102	Sequence 102, App
58	34	58.6	15	1	US-08-465-325-100	Sequence 100, App
59	34	58.6	15	1	US-08-465-325-101	Sequence 101, App
60	34	58.6	15	1	US-08-465-325-116	Sequence 116, App
61	34	58.6	15	4	US-09-115-737-100	Sequence 100, App
62	34	58.6	15	4	US-09-115-737-101	Sequence 101, App
63	34	58.6	15	4	US-09-115-737-116	Sequence 116, App
64	34	58.6	15	5	PCT-US91-05047-66	Sequence 66, App1
65	34	58.6	16	1	US-08-193-521-6	Sequence 6, App1
66	34	58.6	16	1	US-08-434-120-100	Sequence 100, App
67	34	58.6	16	1	US-08-465-325-99	Sequence 99, App1
68	34	58.6	16	4	US-09-115-737-99	Sequence 99, App1
69	34	58.6	17	1	US-07-725-331-47	Sequence 47, App1
70	34	58.6	17	1	US-07-725-331-48	Sequence 48, App1
71	34	58.6	17	1	US-07-725-331-54	Sequence 54, App1
72	34	58.6	17	1	US-07-725-331-55	Sequence 55, App1
73	34	58.6	17	1	US-07-725-331-56	Sequence 56, App1
74	34	58.6	17	2	US-08-569-188-7	Sequence 7, App1
75	34	58.6	17	5	PCT-US91-05047-46	Sequence 46, App1
76	34	58.6	17	5	PCT-US91-05047-47	Sequence 47, App1
77	34	58.6	17	5	PCT-US91-05047-48	Sequence 48, App1
78	34	58.6	17	5	PCT-US91-05047-54	Sequence 54, App1
79	34	58.6	17	5	PCT-US91-05047-55	Sequence 55, App1
80	34	58.6	17	5	PCT-US91-05047-56	Sequence 56, App1
81	34	58.6	17	5	PCT-US91-05047-57	Sequence 57, App1
82	34	58.6	18	1	US-07-725-331-3	Sequence 3, App1
83	34	58.6	18	1	US-07-725-331-5	Sequence 5, App1
84	34	58.6	18	1	US-07-725-331-6	Sequence 6, App1
85	34	58.6	18	1	US-07-725-331-7	Sequence 7, App1
86	34	58.6	18	1	US-07-725-331-11	Sequence 11, App1
87	34	58.6	18	1	US-07-725-331-11	Sequence 11, App1
88	34	58.6	18	1	US-07-725-331-20	Sequence 20, App1
89	34	58.6	18	1	US-07-725-331-21	Sequence 21, App1
90	34	58.6	18	1	US-07-725-331-22	Sequence 22, App1
91	34	58.6	18	1	US-07-725-331-23	Sequence 23, App1
92	34	58.6	18	1	US-07-725-331-24	Sequence 24, App1
93	34	58.6	18	1	US-07-725-331-30	Sequence 30, App1
94	34	58.6	18	1	US-08-295-085-1	Sequence 1, App1
95	34	58.6	18	1	US-08-295-085-9	Sequence 9, App1
96	34	58.6	18	1	US-08-944-133-53	Sequence 53, App1
97	34	58.6	18	1	PCT-US91-05047-3	Sequence 3, App1
98	34	58.6	18	5	PCT-US91-05047-5	Sequence 5, App1
99	34	58.6	18	5		
100	34	58.6	18	5		

101	34	58.6	18	5	PCT-US91-05047-6	Sequence 6, Appli	174	33	56.9	387	4	US-09-134-001C-5036	Sequence 5036, Ap
102	34	58.6	18	5	PCT-US91-05047-7	Sequence 7, Appli	175	33	56.9	445	4	US-08-083-945C-7	Sequence 2, Appli
103	34	58.6	18	5	PCT-US91-05047-11	Sequence 11, Appli	176	33	56.9	445	4	US-08-083-945C-7	Sequence 7, Appli
104	34	58.6	18	5	PCT-US91-05047-19	Sequence 19, Appli	177	33	56.9	492	3	US-08-724-46B-2	Sequence 2, Appli
105	34	58.6	18	5	PCT-US91-05047-20	Sequence 20, Appli	178	33	56.9	492	3	US-08-882-164D-2	Sequence 2, Appli
106	34	58.6	18	5	PCT-US91-05047-21	Sequence 21, Appli	179	33	56.9	539	2	US-08-735-041A-2	Sequence 2, Appli
107	34	58.6	18	5	PCT-US91-05047-22	Sequence 22, Appli	180	33	56.9	539	3	US-09-190-476B-2	Sequence 2, Appli
108	34	58.6	18	5	PCT-US91-05047-23	Sequence 23, Appli	181	33	56.9	539	3	US-09-190-889A-2	Sequence 2, Appli
109	34	58.6	18	5	PCT-US91-05047-24	Sequence 24, Appli	182	33	56.9	539	3	US-09-190-889B-2	Sequence 2, Appli
110	34	58.6	18	5	PCT-US91-05047-30	Sequence 30, Appli	183	33	56.9	539	5	PCT-US95-09261-2	Sequence 2, Appli
111	34	58.6	18	5	PCT-US91-05047-31	Sequence 31, Appli	184	33	56.9	539	5	US-08-735-041A-2	Sequence 4, Appli
112	34	58.6	18	5	PCT-US95-10741-1	Sequence 1, Appli	185	33	56.9	604	2	US-09-190-476B-4	Sequence 4, Appli
113	34	58.6	18	5	PCT-US95-10741-9	Sequence 9, Appli	186	33	56.9	604	3	US-09-190-889A-4	Sequence 4, Appli
114	34	58.6	22	1	US-07-725-331-60	Sequence 60, Appli	187	33	56.9	604	4	US-09-190-938B-4	Sequence 4, Appli
115	34	58.6	22	3	US-08-940-095-54	Sequence 54, Appli	188	33	56.9	614	4	US-09-134-001C-3061	Sequence 3061, Ap
116	34	58.6	22	3	US-08-940-093-54	Sequence 54, Appli	189	33	56.9	674	4	US-08-961-083-200	Sequence 200, Ap
117	34	58.6	22	3	US-08-940-096-54	Sequence 54, Appli	190	33	56.9	676	1	US-08-343-785-2	Sequence 2, Appli
118	34	58.6	22	4	US-09-465-719-54	Sequence 54, Appli	191	33	56.9	676	1	US-08-343-785-8	Sequence 2, Appli
119	34	58.6	22	4	US-09-453-605-54	Sequence 54, Appli	192	33	56.9	676	2	US-08-462-221-2	Sequence 2, Appli
120	34	58.6	22	4	US-09-453-838-54	Sequence 54, Appli	193	33	56.9	676	2	US-08-462-221-8	Sequence 2, Appli
121	34	58.6	22	5	PCT-US91-05047-60	Sequence 60, Appli	194	33	56.9	676	3	US-08-946-458-2	Sequence 2, Appli
122	34	58.6	26	1	US-07-725-331-61	Sequence 61, Appli	195	33	56.9	676	3	US-08-946-458-8	Sequence 2, Appli
123	34	58.6	26	5	PCT-US91-05047-61	Sequence 61, Appli	196	33	56.9	676	4	US-09-708-426-10	Sequence 10, Appli
124	34	58.6	30	1	US-07-725-331-62	Sequence 62, Appli	197	33	56.9	701	2	US-08-735-041A-6	Sequence 6, Appli
125	34	58.6	30	1	PCT-US91-05047-62	Sequence 62, Appli	198	33	56.9	701	3	US-09-190-476B-6	Sequence 6, Appli
126	34	58.6	36	1	US-07-725-331-63	Sequence 63, Appli	199	33	56.9	701	3	US-09-190-889A-6	Sequence 6, Appli
127	34	58.6	36	5	PCT-US91-05047-63	Sequence 63, Appli	200	33	56.9	701	4	US-09-190-889B-6	Sequence 6, Appli
128	34	58.6	80	4	US-09-048-502-11	Sequence 10, Appli	201	33	56.9	855	2	US-08-468-558-2	Sequence 2, Appli
129	34	58.6	80	4	US-09-048-502-11	Sequence 11, Appli	202	33	56.9	855	2	US-08-676-444-2	Sequence 2, Appli
130	34	58.6	89	4	US-09-048-502-2	Sequence 2, Appli	203	33	56.9	855	4	US-08-697-954-2	Sequence 2, Appli
131	34	58.6	89	4	US-09-048-502-3	Sequence 3, Appli	204	33	56.9	1070	4	US-08-697-954-4	Sequence 2, Appli
132	34	58.6	102	4	US-09-048-502-1	Sequence 1, Appli	205	33	56.9	1205	4	US-09-330-330-1	Sequence 1, Appli
133	34	58.6	110	1	US-08-434-705B-15	Sequence 15, Appli	206	32	55.2	2154	2	US-08-841-349-4	Sequence 4, Appli
134	34	58.6	110	2	US-09-086-201-15	Sequence 15, Appli	207	32	55.2	11	1	US-08-465-325-141	Sequence 141, App
135	34	58.6	116	1	US-08-434-705B-8	Sequence 6, Appli	208	32	55.2	14	1	US-07-725-331-1	Sequence 1, Appli
136	34	58.6	116	1	US-08-434-705B-8	Sequence 6, Appli	209	32	55.2	14	5	PCT-US91-05047-1	Sequence 1, Appli
137	34	58.6	116	2	US-09-086-201-6	Sequence 6, Appli	210	32	55.2	18	1	US-07-725-331-9	Sequence 9, Appli
138	34	58.6	124	2	US-09-086-201-8	Sequence 8, Appli	211	32	55.2	18	1	US-07-725-331-10	Sequence 10, Appli
139	34	58.6	124	2	US-08-434-705B-16	Sequence 16, Appli	212	32	55.2	18	1	US-07-725-331-18	Sequence 18, Appli
140	34	58.6	129	1	US-09-086-201-16	Sequence 16, Appli	213	32	55.2	18	1	US-07-725-331-29	Sequence 29, Appli
141	34	58.6	129	1	US-08-434-705B-4	Sequence 4, Appli	214	32	55.2	18	5	PCT-US91-05047-9	Sequence 9, Appli
142	34	58.6	129	2	US-09-086-201-4	Sequence 4, Appli	215	32	55.2	18	5	PCT-US91-05047-10	Sequence 10, Appli
143	34	58.6	184	3	US-08-741-411-1	Sequence 1, Appli	216	32	55.2	18	5	PCT-US91-05047-18	Sequence 18, Appli
144	34	58.6	327	1	US-08-080-386-2	Sequence 2, Appli	217	32	55.2	18	5	PCT-US91-05047-29	Sequence 29, Appli
145	34	58.6	327	2	US-08-390-000A-2	Sequence 2, Appli	218	32	55.2	29	2	US-08-464-182A-25	Sequence 25, Appli
146	34	58.6	348	3	US-08-869-696-4	Sequence 4, Appli	219	32	55.2	29	2	US-08-693-274A-15	Sequence 15, Appli
147	34	58.6	395	3	US-08-869-696-3	Sequence 3, Appli	220	32	55.2	29	2	US-08-406-271-25	Sequence 25, Appli
148	34	58.6	465	3	US-08-948-997-6	Sequence 6, Appli	221	32	55.2	107	4	US-09-228-986-125	Sequence 125, App
149	34	58.6	465	4	US-09-348-817A-6	Sequence 6, Appli	222	32	55.2	126	4	US-09-284-033-5	Sequence 5, Appli
150	34	58.6	497	3	US-08-869-696-9	Sequence 9, Appli	223	32	55.2	126	4	US-08-729-834B-5	Sequence 5, Appli
151	34	58.6	503	4	US-08-740-223A-18	Sequence 18, Appli	224	32	55.2	161	1	US-08-948-195-3	Sequence 3, Appli
152	34	58.6	503	4	US-09-202-491-4	Sequence 4, Appli	225	32	55.2	161	2	US-09-094-080-3	Sequence 3, Appli
153	34	58.6	503	4	US-09-202-491-10	Sequence 10, Appli	226	32	55.2	161	2	US-09-216-655-3	Sequence 23, Appli
154	34	58.6	503	4	US-09-709-188-18	Sequence 18, Appli	227	32	55.2	177	4	US-09-522-433B-23	Sequence 5, Appli
155	34	58.6	556	3	US-08-869-696-2	Sequence 2, Appli	228	32	55.2	190	3	US-08-741-411-5	Sequence 11, Appli
156	34	58.6	586	1	US-08-198-446B-19	Sequence 19, Appli	229	32	55.2	214	3	US-09-004-731-13	Sequence 13, Appli
157	34	58.6	586	2	US-08-870-693-19	Sequence 2, Appli	230	32	55.2	225	4	US-08-749-699-13	Sequence 13, Appli
158	34	58.6	934	1	US-08-457-176-2	Sequence 2, Appli	231	32	55.2	225	4	US-09-004-729-13	Sequence 13, Appli
159	34	58.6	934	1	US-08-457-175-2	Sequence 2, Appli	232	32	55.2	225	4	US-09-004-729-13	Sequence 13, Appli
160	34	58.6	934	4	US-08-709-784-1	Sequence 1, Appli	233	32	55.2	228	4	US-08-749-699-10	Sequence 10, Appli
161	34	58.6	934	4	US-09-651-656-3	Sequence 3, Appli	234	32	55.2	228	4	US-09-004-729-10	Sequence 10, Appli
162	34	58.6	934	4	US-09-650-855-3	Sequence 3, Appli	235	32	55.2	228	4	US-09-004-729-10	Sequence 10, Appli
163	34	58.6	1001	4	US-09-060-410-2	Sequence 2, Appli	236	32	55.2	231	2	US-08-557-309B-38	Sequence 38, Appli
164	34	58.6	4536	1	US-09-180-422B-27	Sequence 27, Appli	237	32	55.2	231	3	US-08-834-306-38	Sequence 38, Appli
165	34	58.6	4536	1	US-07-725-331-26	Sequence 26, Appli	238	32	55.2	231	4	US-08-993-674A-38	Sequence 38, Appli
166	33	56.9	18	3	US-08-940-095-205	Sequence 205, App	239	32	55.2	231	4	US-09-256-976-38	Sequence 38, Appli
167	33	56.9	18	3	US-08-940-093-205	Sequence 205, App	240	32	55.2	277	2	US-08-568-459A-15	Sequence 15, Appli
168	33	56.9	18	3	US-08-940-096-205	Sequence 205, App	241	32	55.2	277	2	US-08-487-826B-27	Sequence 27, Appli
169	33	56.9	18	4	US-09-465-719-205	Sequence 205, App	242	32	55.2	277	4	US-09-210-288-15	Sequence 15, Appli
170	33	56.9	18	4	US-09-453-605-205	Sequence 205, App	243	32	55.2	283	1	US-08-723-886-3	Sequence 3, Appli
171	33	56.9	18	4	US-09-453-838-205	Sequence 26, Appli	244	32	55.2	283	3	US-08-972-902-8	Sequence 8, Appli
172	33	56.9	18	5	PCT-US91-05047-26	Sequence 26, Appli	245	32	55.2	283	3	US-08-972-902-8	Sequence 9, Appli
173	33	56.9	218	4	US-09-615-192A-269	Sequence 269, App	246	32	55.2	283	4	US-09-466-257A-8	Sequence 8, Appli

247	32	55.2	283	4	US-09-466-257A-10	Sequence 10, Appl
248	32	55.2	283	4	US-08-424-797A-3	Sequence 3, Appl
249	32	55.2	306	1	US-08-312-387B-4	Sequence 4, Appl
250	32	55.2	306	1	US-08-683-426-4	Sequence 4, Appl
251	32	55.2	306	1	US-08-683-458-4	Sequence 4, Appl
252	32	55.2	306	1	US-08-878-360-4	Sequence 4, Appl
253	32	55.2	306	4	US-08-478-140B-4	Sequence 4, Appl
254	32	55.2	306	4	US-09-333-412-4	Sequence 4, Appl
255	32	55.2	306	4	US-09-338-943-4	Sequence 4, Appl
256	32	55.2	316	4	US-08-729-594A-36	Sequence 36, Appl
257	32	55.2	316	4	US-08-937-993-36	Sequence 36, Appl
258	32	55.2	317	2	US-09-066-075-2	Sequence 2, Appl
259	32	55.2	317	2	US-08-518-615A-2	Sequence 2, Appl
260	32	55.2	317	3	US-08-951-889-2	Sequence 2, Appl
261	32	55.2	317	4	US-09-472-857-2	Sequence 2, Appl
262	32	55.2	394	4	US-09-202-893B-2	Sequence 2, Appl
263	32	55.2	399	4	US-08-284-033-2	Sequence 2, Appl
264	32	55.2	399	4	US-08-729-834B-2	Sequence 2, Appl
265	32	55.2	473	5	US-08-272-255-11	Sequence 11, Appl
266	32	55.2	473	5	PCT-US95-08565-11	Sequence 11, Appl
267	32	55.2	513	4	US-09-134-001C-4386	Sequence 4386, Ap
268	32	55.2	513	4	US-09-134-001C-4629	Sequence 4629, Ap
269	32	55.2	515	4	US-08-961-083-22	Sequence 22, Appl
270	32	55.2	527	4	US-08-426-509A-10	Sequence 10, Appl
271	32	55.2	527	5	PCT-US95-05008-10	Sequence 10, Appl
272	32	55.2	599	1	US-08-391-615-4	Sequence 4, Appl
273	32	55.2	968	1	US-08-426-236-2	Sequence 2, Appl
274	32	55.2	1365	6	US-09-376-330-18	Sequence 18, Appl
275	32	55.2	1365	6	5194600-4	Patent No. 5194600
276	32	55.2	2291	2	US-08-286-819A-29	Sequence 29, Appl
277	32	55.2	2291	3	US-08-980-357-29	Sequence 29, Appl
278	32	55.2	2293	4	US-09-368-590-2	Sequence 2, Appl
279	32	55.2	3174	2	US-08-477-451-3	Sequence 3, Appl
280	32	55.4	11	1	US-08-465-325-145	Sequence 146, App
281	31	53.4	11	4	US-07-725-331-49	Sequence 146, App
282	31	53.4	17	1	PCT-US91-05047-49	Sequence 49, Appl
283	31	53.4	17	5	5304631-10	Patent No. 5304631
284	31	53.4	19	6	US-08-940-095-39	Sequence 39, Appl
285	31	53.4	22	3	US-08-940-095-184	Sequence 184, App
286	31	53.4	22	3	US-08-940-095-184	Sequence 184, App
287	31	53.4	22	3	US-08-940-093-39	Sequence 39, Appl
288	31	53.4	22	3	US-08-940-093-39	Sequence 39, Appl
289	31	53.4	22	3	US-08-940-093-164	Sequence 164, App
290	31	53.4	22	3	US-08-940-093-184	Sequence 184, App
291	31	53.4	22	3	US-08-940-096-39	Sequence 39, Appl
292	31	53.4	22	3	US-08-940-096-164	Sequence 164, App
293	31	53.4	22	3	US-08-940-096-184	Sequence 184, App
294	31	53.4	22	4	US-09-465-719-39	Sequence 39, Appl
295	31	53.4	22	4	US-09-465-719-164	Sequence 164, App
296	31	53.4	22	4	US-09-465-719-184	Sequence 184, App
297	31	53.4	22	4	US-09-453-605-39	Sequence 39, Appl
298	31	53.4	22	4	US-09-453-605-164	Sequence 164, App
299	31	53.4	22	4	US-09-453-605-184	Sequence 184, App
300	31	53.4	22	4	US-09-453-838-39	Sequence 39, Appl

ALIGNMENTS

RESULT 1
 US-08-046-431A-2
 : Sequence 2, Application US/08046431A
 : Patent No. 5420252
 :
 : GENERAL INFORMATION:
 : APPLICANT: KATO, HIROYUKI
 : APPLICANT: YOSHITAKE, SHINJI
 : APPLICANT: SUZUKI, SUGURU
 : APPLICANT: SUZUKI, NOBORU
 : APPLICANT: SETO, TOSHIO
 : APPLICANT: NAGAOKA, NAOKO
 : APPLICANT: MIZU, YOSHIHARU
 : TITLE OF INVENTION: HUMAN ANTIITHROMBIN III MUTANT
 : NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:
 ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS, P.C.
 STREET: 2026 Rambling Road
 CITY: Kalamazoo
 STATE: Michigan
 COUNTRY: USA
 ZIP: 49008-1699
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
 COMPUTER: IBM PC/XT/AT Compatible
 OPERATING SYSTEM: MS-DOS 5.0
 SOFTWARE: WordPerfect 5.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/046,431A
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP90488
 FILING DATE: 10-APR-1992
 APPLICATION NUMBER: JP31855
 FILING DATE: 22-FEB-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Terrance F. Chapman
 REGISTRATION NUMBER: 32549
 REFERENCE/DOCKET NUMBER: Furuya Case 1286
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (616) 381-1156
 TELEFAX: (616) 381-5465
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 464 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 DESCRIPTION: protein
 US-08-046-431A-2

Query Match 69.0%; Score 40; DB 1; Length 464;
 Best Local Similarity 63.6%; Pred. No. 1.8e+02;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 LQEWLKLKKM 11
 Db 336 LQEWLDELEEM 346

RESULT 2
 US-08-138-641-2
 : Sequence 2, Application US/08138641
 : Patent No. 5474921
 :
 : GENERAL INFORMATION:
 : APPLICANT: Kodian, Kenneth S.
 : APPLICANT: Pompiano, David L.
 : TITLE OF INVENTION: ASSAY TO DETERMINE INHIBITORS OF
 : NUMBER OF SEQUENCES: 9
 : CORRESPONDENCE ADDRESS:
 ADDRESSEE: David A. Muthard
 STREET: P.O. Box 2000, 126 E. Lincoln Avenue
 CITY: Rahway
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07065
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/138,641
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:

NAME: Muthard, David A.
REGISTRATION NUMBER: 35,297
REFERENCE/DOCKET NUMBER: 18937
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3903
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1290 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-138-641-2

Query Match 67.2%; Score 39; DB 1; Length 1290;
Best Local Similarity 45.5%; Pred. No. 5.8e+02;
Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOEWLKKLKKM 11
||:||||:|
Db 919 LQDWKRIREV 929

RESULT 3
US-08-138-133-2
Sequence 2, Application US/08138133
Patent No. 5519163
GENERAL INFORMATION:
APPLICANT: GIBBS, JACKSON B.
APPLICANT: KOBLAN, KENNETH S.
APPLICANT: MACLEOD, ANGUS M.
APPLICANT: MERCHANT, KEVIN J.
TITLE OF INVENTION: INHIBITORS OF PHOSPHOINOSITIDE-SPECIFIC
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID A. MUTHARD
STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
CITY: RAHWAY
STATE: NEW JERSEY
COUNTRY: U.S.A.
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/138,133
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: MUTHARD, DAVID A.
REGISTRATION NUMBER: 35,297
REFERENCE/DOCKET NUMBER: 18938
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3903
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1290 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-138-133-2

Query Match 67.2%; Score 39; DB 1; Length 1290;
Best Local Similarity 45.5%; Pred. No. 5.8e+02;
Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
OY 1 LOEWLKKLKKM 11

||:||||:|
Db 919 LQDWKRIREV 929

RESULT 4
US-08-270-076A-11
Sequence 11, Application US/08270076A
Patent No. 5667986
GENERAL INFORMATION:
APPLICANT: Sleep, Darrell
APPLICANT: Goodey, Andrew R
APPLICANT: Vakerlia, Diana
TITLE OF INVENTION: Yeast Promoter
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: The BOC Group, Inc.
STREET: 100 Mountain Avenue, Murray Hill
CITY: New Providence
STATE: New Jersey
COUNTRY: USA
ZIP: 07974
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/270,076A
FILING DATE: 01-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8923521.2
FILING DATE: 18-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/597,687
FILING DATE: 16-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/925,286
FILING DATE: 04-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R. Hain
REGISTRATION NUMBER: 24864
REFERENCE/DOCKET NUMBER: 92H834-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908/771-6292
TELEFAX: 908/771-6159
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 806 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-270-076A-11

Query Match 65.5%; Score 38; DB 1; Length 806;
Best Local Similarity 54.5%; Pred. No. 5.3e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKKLKKM 11
||:||||:|
Db 338 LEWLOROKKV 348

RESULT 5
US-08-411-768B-4
Sequence 4, Application US/08411768B
Patent No. 6083712
GENERAL INFORMATION:
APPLICANT: Oihen Birch
APPLICANT: Johann Brass
APPLICANT: Martin Fuhrmann
APPLICANT: Nicholas Shaw
TITLE OF INVENTION: Biotechnological Method

```

; TITLE OF INVENTION: of Producing Biotin
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect
; SOFTWARE: Version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/411,768B
; FILING DATE: 31-March-95
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 3124/92
; FILING DATE: 02-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 2134/93
; FILING DATE: 15-JUL-1993
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 429 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-411-768B-4

Query Match
Best Local Similarity 55.6%; Score 37; DB 3; Length 429;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EWLKRLKRM 11
DB 226 EWLKRLKRI 234

RESULT 6
US-08-401-068-4
; Sequence 4, Application US/08401068
; Patent No. 5859335
; GENERAL INFORMATION:
; APPLICANT: Patton, David
; TITLE OF INVENTION: Enhanced Biotin Biosynthesis in Plant Tissue
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/401,068
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/351,970
; FILING DATE: 08-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-401-068-4

Query Match
Best Local Similarity 55.6%; Score 37; DB 2; Length 431;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EWLKRLKRM 11
DB 228 EWLKRLKRI 236

RESULT 7
US-08-846-338-4
; Sequence 4, Application US/08846338
; Patent No. 5869719
; GENERAL INFORMATION:
; APPLICANT: Patton, David
; TITLE OF INVENTION: Transgenic Plants Having Increased Biotin Content
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5869719artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/846,338
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-846-338-4

Query Match
Best Local Similarity 55.6%; Score 37; DB 2; Length 431;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EWLKRLKRM 11
DB 228 EWLKRLKRI 236

RESULT 8
US-09-475-316A-118
; Sequence 118, Application US/09475316A
; Patent No. 6210942
; GENERAL INFORMATION:
; APPLICANT: Lewis, No. 6210942man G.
; APPLICANT: Davin, Laurence B.

```

APPLICANT: Dinkova-Kostova, Albena T.
APPLICANT: Fujita, Masayuki
APPLICANT: Gang, David R.
APPLICANT: Saikaneu, Simo
APPLICANT: Ford, Joshua D
TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASES,
TITLE OF INVENTION: RECOMBINANT DIRIGENT PROTEINS AND METHODS OF USE
FILE REFERENCE: W09-1-13793
CURRENT APPLICATION NUMBER: US/09/475,316A
CURRENT FILING DATE: 1999-12-30
PRIOR APPLICATION NUMBER: 09/307,653
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: PCT/US97/20391
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/054,380
PRIOR FILING DATE: 1997-07-31
PRIOR APPLICATION NUMBER: 60/030,522
PRIOR FILING DATE: 1996-11-08
NUMBER OF SEQ ID NOS: 122
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 118
LENGTH: 314
TYPE: PRT
ORGANISM: Schisandra chinensis
US-09-475-316A-118

Query Match 62.1%; Score 36; DB 4; Length 314;
Best Local Similarity 60.0%; Pred. No. 4.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 OEMLKKKK 11
DB 254 OEMLKMEGM 263

RESULT 9
US-09-134-001C-3153
Sequence 3153, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO: 3153
LENGTH: 560
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3153

Query Match 62.1%; Score 36; DB 4; Length 560;
Best Local Similarity 55.6%; Pred. No. 7.2e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOEWLKLK 9
DB 244 IOEWLQEMK 252

RESULT 10
US-08-227-536-2
Sequence 2, Application US/08227536
Patent No. 5658784
GENERAL INFORMATION:
APPLICANT: Eckner, Richard
ATTORNEY/AGENT INFORMATION:

APPLICANT: Livingston, David
TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION
TITLE OF INVENTION: FACTOR P300 AND USES OF P300
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,536
FILING DATE: 14-APR-1994
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Williams Ph.D., Kathleen A.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: DECI-308XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2414 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-227-536-2

Query Match 62.1%; Score 36; DB 1; Length 2414;
Best Local Similarity 75.0%; Pred. No. 2.5e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKLK 8
DB 1463 LOEWYKKM 1470

RESULT 11
PCT-US95-04682-2
Sequence 2, Application PC/TUS9504682
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
TITLE OF INVENTION: FACTOR P300 AND USES OF P300
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,536
FILING DATE: 14-APR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

NAME: Holliday C. Heine, Ph.D.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DCCI-308Xq999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2280
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2414 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-04682-2

Query Match 62.1%; Score 36; DB 5; Length 2414;
Best Local Similarity 75.0%; Pred. No. 2.5e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKL 8
DB 1463 LOEWYKKM 1470

RESULT 12
US-08-194-468-2
Sequence 2, Application US/08194468
Patent No. 5750336
GENERAL INFORMATION:
APPLICANT: Montminy, Marc R.
TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF
TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND MITOGEN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/194,468
FILING DATE: 10-FEB-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9672
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)-546-4737
TELEFAX: (619)-546-9392
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2441 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-194-468-2

Query Match 62.1%; Score 36; DB 1; Length 2441;
Best Local Similarity 75.0%; Pred. No. 2.5e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKL 8
DB 1500 LOEWYKKM 1507

RESULT 13
US-08-961-739-2
Sequence 2, Application US/08961739A
Patent No. 6063583
GENERAL INFORMATION:
APPLICANT: Montminy, Marc R.
TITLE OF INVENTION: Methods for Treating Diabetes Mellitus
FILE REFERENCE: SALK1650-1
CURRENT APPLICATION NUMBER: US/08/961,739A
CURRENT FILING DATE: 1997-10-31
EARLIER APPLICATION NUMBER: US 194,468
EARLIER FILING DATE: 1994-02-10
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 2441
TYPE: PRT
ORGANISM: Mus
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)-(2441)
OTHER INFORMATION: Xaa - Any Amino Acid
US-08-961-739-2

Query Match 62.1%; Score 36; DB 3; Length 2441;
Best Local Similarity 75.0%; Pred. No. 2.5e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKL 8
DB 1500 LOEWYKKM 1507

RESULT 14
US-09-514-247A-8
Sequence 8, Application US/09514247A
Patent No. 6365361
GENERAL INFORMATION:
APPLICANT: TANIGUCHI, Tomoyasu
APPLICANT: TANIGUCHI, Tomoyasu
APPLICANT: MIZUKAMI, Junko
TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO
FILE REFERENCE: TANIGUCHI-6
CURRENT APPLICATION NUMBER: US/09/514,247A
CURRENT FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: PCT/JP98/03734
PRIOR FILING DATE: 1998-08-24
PRIOR APPLICATION NUMBER: JP231084/1997
PRIOR FILING DATE: 1997-08-27
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin version 3.0
SEQ ID NO 8
LENGTH: 2441
TYPE: PRT
ORGANISM: mouse
US-09-514-247A-8

Query Match 62.1%; Score 36; DB 4; Length 2441;
Best Local Similarity 75.0%; Pred. No. 2.5e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKL 8
DB 1500 LOEWYKKM 1507

RESULT 15
US-09-514-247A-10
Sequence 10, Application US/09514247A
Patent No. 6365361
GENERAL INFORMATION:
APPLICANT: TANABE SEIYAKU CO. LTD.
APPLICANT: TANIGUCHI, Tomoyasu

APPLICANT: MIZUKAMI, Junko
TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PPA
FILE REFERENCE: TANIGUCHI-6
CURRENT APPLICATION NUMBER: US/09/514,247A
CURRENT FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: PCT/JP98/03734
PRIOR FILING DATE: 1998-08-24
PRIOR APPLICATION NUMBER: JP231084/1997
PRIOR FILING DATE: 1997-08-27
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.0
SEQ ID NO 10
LENGTH: 2442
TYPE: PRT
ORGANISM: human
US-09-514-247A-10

Query Match 62.1%, Score 36; DB 4; Length 2442;
Best Local Similarity 75.0%; Pred. No. 2.5e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LQEWLKKL 8
DB 1499 LQEWYKKM 1506

RESULT 16
US-07-789-179-2
Sequence 2, Application US/07789179
Patent No. 5545727
GENERAL INFORMATION:
APPLICANT: HOFFMAN, STEPHEN J.
APPLICANT: LOOKER, DOUGLAS L.
APPLICANT: ROSENDAHL, MARY S.
APPLICANT: STETLER, GARY L.
APPLICANT: WAGENBACH, MICHAEL
APPLICANT: ANDERSON, DAVID C.
APPLICANT: MATHEWS, ANTONY J.
TITLE OF INVENTION: PRODUCTION AND USE OF HEMOGLOBIN AND
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 SEVENTH STREET, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/789,179
FILING DATE: 19911108
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/671,707
FILING DATE: 01-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/374,161
FILING DATE: 30-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/379,116
FILING DATE: 13-JUL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/349,623
FILING DATE: 10-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, IVER P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: HOFFMAN-5B

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-789-179-2

Query Match 60.3%, Score 35; DB 1; Length 16;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LQEWLKKLKKM 11
DB 3 LEELLKKLKL 13

RESULT 17
US-08-240-712-9
Sequence 9, Application US/08240712
Patent No. 5599907
GENERAL INFORMATION:
APPLICANT: ANDERSON, DAVID C.
APPLICANT: MATHEWS, ANTONY JAMES
APPLICANT: STETLER, GARY L.
TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/240,712
FILING DATE: 09-MAY-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/09752
FILING DATE: 13-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, IVER P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: ANDERSON-6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-240-712-9

Query Match 60.3%, Score 35; DB 1; Length 16;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY	1	LQEWLKKLKM	11
		: :	
Db	3	LEELLKKLKE	13

RESULT 18
US-08-443-890-9
; Sequence 9, Application US/08443890

1 GENERAL INFORMATION:
2 APPLICANT: ANDERSON, DAVID C.
3 APPLICANT: MATHEWS, ANTHONY JAMES
4 APPLICANT: STEETLER, GARY L.
5 TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC
6 TITLE OF INVENTION: HEMOGLOBINS
7 NUMBER OF SEQUENCES: 35
8 CORRESPONDENCE ADDRESS:
9 ADDRESSEE: Brody and Netmark
10 STREET: 419 Seventh Street, N.W., Suite 300
11 CITY: Washington
12 STATE: D.C.
13 COUNTRY: USA
14 ZIP: 20004

Query Match	60.3%	Score 35:	DB 1:	Length 16:
Best Local Similarity	63.6%	Pred. No. 49:		
Matches 7:	Conservative 3:	Mismatches 1:	Indels 0:	Gaps 0:

RESULT 19
US-08-444-942-2
Sequence 2, Application US/08444942-2
Patent No. 5744329
GENERAL INFORMATION:
APPLICANT: HOFFMAN, STEPHEN J.
APPLICANT: LOOKER, DOUGLAS L.
APPLICANT: ROSENDAHL, MARY S.
APPLICANT: STETLER, GARY L.
APPLICANT: WAGENBACH, MICHAEL

1 APPLICANT: ANDERSON, DAVID C.
2 APPLICANT: MATHEWS, ANTONY J.
3 TITLE OF INVENTION: PRODUCTION AND USE OF HEMOGLOBIN AND
4 TITLE OF INVENTION: PRODUCTION AND USE OF HEMOGLOBIN AND
5 TITLE OF INVENTION: PRODUCTION AND USE OF HEMOGLOBIN AND
6 NUMBER OF SEQUENCES: 7
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: BROWDY AND NETMARK
9 STREET: 419 SEVENTH STREET, N.W.
10 CITY: WASHINGTON
11 STATE: D.C.
12 COUNTRY: U.S.A.
13 ZIR: 20004

Query Match	60.3%	Score 35:	DB 1:	Length 16:
Best Local Similarity	63.6%	Pred. No.	49:	
Matches 7, Conservative	3:	Mismatches	1:	Indels 0:
				Gaps 0:

RESULT 20
US-08-446-105-2
Sequence 2, Application US/08446105-2
Patent No. 5796227
GENERAL INFORMATION:
APPLICANT: HOFFMAN, STEPHEN J.
APPLICANT: LOOKER, DOUGLAS L.
APPLICANT: ROSENDAHL, MARY S.
APPLICANT: STETLER, GARY L.
APPLICANT: WAGENBICH, MICHAEL
APPLICANT: ANDERSON, DAVID C.
APPLICANT: MATTHEWS, ANTONY J.

;; TITLE OF INVENTION: PRODUCTION AND USE OF HEMOGLOBIN AND
;; TITLE OF INVENTION: ANALOGUES THEREOF
;; NUMBER OF SEQUENCES: 7
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: BROWDY AND NEIMARK
;; STREET: 419 SEVENTH STREET, N.W.
;; CITY: WASHINGTON
;; STATE: D.C.
;; COUNTRY: U.S.A.
;; ZIP: 20004
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/446,105
;; FILING DATE:
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/789,179
;; FILING DATE: 08-NOV-1991
;; APPLICATION NUMBER: US/07/671,707
;; FILING DATE: 01-APR-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/374,161
;; FILING DATE: 30-JUN-1989
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/379,116
;; FILING DATE: 13-JUL-1989
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/349,623
;; FILING DATE: 10-MAY-1989
;; ATTORNEY/AGENT INFORMATION:
;; NAME: COOPER, IYER P.
;; REGISTRATION NUMBER: 28,005
;; REFERENCE/DOCKET NUMBER: HOFFMAN-5B
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-628-5197
;; TELEFAX: 202-737-3528
;; TELEX: 248633
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 16 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;;
US-08-446-105-2
;;
Query Match 60.3%; Score 35; DB 1; Length 16;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LQEWLKKIKKM 11
Db 3 LEEILKKIKLKL 13

RESULT 21
US-08-444-939-2
;; Sequence 2, Application US/08444939
;; Patent No. 5801019
;; GENERAL INFORMATION:
;; APPLICANT: HOFFMAN, STEPHEN J.
;; APPLICANT: LOOKER, DOUGLAS L.
;; APPLICANT: ROSENDALH, MARY S.
;; APPLICANT: STETLER, GARY L.
;; APPLICANT: WAGENBACH, MICHAEL
;; APPLICANT: ANDERSON, DAVID C.
;; APPLICANT: MATHEWS, ANTONY J.
;; TITLE OF INVENTION: PRODUCTION AND USE OF HEMOGLOBIN AND
;; ANALOGUES THEREOF

;; NUMBER OF SEQUENCES: 7
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: BROWDY AND NEIMARK
;; STREET: 419 SEVENTH STREET, N.W.
;; CITY: WASHINGTON
;; STATE: D.C.
;; COUNTRY: U.S.A.
;; ZIP: 20004
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/444,939
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/789,179
;; FILING DATE: 08-NOV-1991
;; APPLICATION NUMBER: US/07/671,707
;; FILING DATE: 01-APR-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/374,161
;; FILING DATE: 30-JUN-1989
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/379,116
;; FILING DATE: 13-JUL-1989
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/349,623
;; FILING DATE: 10-MAY-1989
;; ATTORNEY/AGENT INFORMATION:
;; NAME: COOPER, IYER P.
;; REGISTRATION NUMBER: 28,005
;; REFERENCE/DOCKET NUMBER: HOFFMAN-5B
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-628-5197
;; TELEFAX: 202-737-3528
;; TELEX: 248633
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 16 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;;
US-08-444-939-2
;;
Query Match 60.3%; Score 35; DB 1; Length 16;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LQEWLKKIKKM 11
Db 3 LEEILKKIKLKL 13

RESULT 22
US-08-444-991-2
;; Sequence 2, Application US/08444991
;; Patent No. 584408
;; GENERAL INFORMATION:
;; APPLICANT: HOFFMAN, STEPHEN J.
;; APPLICANT: LOOKER, DOUGLAS L.
;; APPLICANT: ROSENDALH, MARY S.
;; APPLICANT: STETLER, GARY L.
;; APPLICANT: WAGENBACH, MICHAEL
;; APPLICANT: ANDERSON, DAVID C.
;; APPLICANT: MATHEWS, ANTONY J.
;; TITLE OF INVENTION: PRODUCTION AND USE OF HEMOGLOBIN AND
;; ANALOGUES THEREOF
;; NUMBER OF SEQUENCES: 7
;; CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 SEVENTH STREET, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,991
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/789,179
FILING DATE: 08-NOV-1991
APPLICATION NUMBER: US/07/671,707
FILING DATE: 01-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/374,161
FILING DATE: 30-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/379,116
FILING DATE: 13-JUL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/349,623
FILING DATE: 10-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, IVER P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: HOFFMAN-5B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-444-991-2

Query Match 60.3%; Score 35; DB 2; Length 16;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 LOEWLKKLKKM 11
|:| |||||:
Db 3 LEELKKLKEEL 13

RESULT 23
US-08-450-733-2
Sequence 2, Application US/08450733
Patent No. 5844089
GENERAL INFORMATION:
APPLICANT: HOFFMAN, STEPHEN J.
APPLICANT: LOOKER, DOUGLAS L.
APPLICANT: ROSENDAHL, MARY S.
APPLICANT: STETLER, GARY L.
APPLICANT: WAGENBACH, MICHAEL
APPLICANT: ANDERSON, DAVID C.
APPLICANT: MATHEWS, ANTONY J.
TITLE OF INVENTION: PRODUCTION AND USE OF HEMOGLOBIN AND
NUMBER OF INVENTION: ANALOGUES THEREOF
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 SEVENTH STREET, N.W.

CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,733
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/789,179
FILING DATE: 08-NOV-1991
APPLICATION NUMBER: US/07/671,707
FILING DATE: 01-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/374,161
FILING DATE: 30-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/379,116
FILING DATE: 13-JUL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/349,623
FILING DATE: 10-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, IVER P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: HOFFMAN-5B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-450-733-2

Query Match 60.3%; Score 35; DB 2; Length 16;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 LOEWLKKLKKM 11
|:| |||||:
Db 3 LEELKKLKEEL 13

RESULT 24
US-09-058-562-9
Sequence 9, Application US/09058562A
Patent No. 6184356
GENERAL INFORMATION:
APPLICANT: Anderson, David C.
APPLICANT: Mathews, Antony James
APPLICANT: Stetler, Gary L.
TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC HEMOGLOBINS
FILE REFERENCE: EYTB 2087
CURRENT APPLICATION NUMBER: US/09/058,562A
PRIOR FILING DATE: 1998-04-13
PRIOR APPLICATION NUMBER: US 08/240,712
PRIOR FILING DATE: 1994-05-09
PRIOR APPLICATION NUMBER: PCT/US92/09752
PRIOR FILING DATE: 1993-05-13
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 9
LENGTH: 16

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: crosslinker
US-09-058-562-9

```

Query Match	60.3%	Score 35	DB 4	Length 16
Best Local Similarity	63.6%	Pred. No. 49		
Matches 7, Conservative	3	Mismatches	1	Indels 0
				Gaps 0

QY	1	LQEWLKKLKKM	11
		: :	
Db	3	LEELLKKLKE	13

RESULT 25
US-08-444-915A-2
; Sequence 2, Application US/08444915A
; Date of Filing: 03/24/2008

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-444-915A-2

```

Query Match	60.3%	Score 35	DB 4	Length 16
Best Local Similarity	63.6%	Pred. No.	49	
Matches 7, Conservative	3	Mismatches	1	Indels 0: Gaps 0:

```
QY      1 LQEWLKKLKM 11
        | : | | | | : :
Db      3 LEELLKKLKL 13
```

RESULT 26
PCT-US92-09752-9
; Sequence 9, Application PC/TUS9209752
; GENERAL INFORMATION:

Query Match	60.3%	Score 35	DB 5	Length 16
Best Local Similarity	63.6%	Pred. No.	49	
Matches 7; Conservative	3	Mismatches	1	Indels 0; Gaps 0

```
QY 1 LQEWLKKLKKM 11
    1:11111::
Db 3 LEELKKLKKEL 13
```

RESULT 27
5304631-1
Patent No. 5304631
APPLICANT: STEWART, JOHN M.; HAHN, KARL W.; KLIS, WIESLAW A
TITLE OF INVENTION: SYNTHETIC HELIXIYME ENZYMES
NUMBER OF SEQUENCES: 16
CURRENT APPLICATION DATA:

GENERAL INFORMATION:
APPLICANT: Dasseux, Jean-Louis
APPLICANT: Sekul, Renate
APPLICANT: Butner, Klaus
APPLICANT: Cornut, Isabelle
APPLICANT: Metz, Gunther
TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
NUMBER OF SEQUENCES: 258
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/940.096
FILING DATE: 29-SEP-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 009196-0005-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6046166e
FEATURE:
NAME/KEY: Other
LOCATION: 1...18
OTHER INFORMATION: N-terminal acetylated and
OTHER INFORMATION: C-terminal amidated
US-08-940-096-215
Query Match 60.3%; Score 35; DB 3; Length 18;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 LOEWLKKLK 9
DB 10 LEWKKKKLK 18
RESULT 31
US-09-465-719-215
Sequence 215, Application US/09465719
Patent No. 6265377
GENERAL INFORMATION:
APPLICANT: Dasseux, Jean-Louis
APPLICANT: Sekul, Renate
APPLICANT: Butner, Klaus
APPLICANT: Cornut, Isabelle
APPLICANT: Metz, Gunther
TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
NUMBER OF SEQUENCES: 258
CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/465,719
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/940.093
FILING DATE: 29-SEP-1997
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 009196-0006-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6265377e
FEATURE:
NAME/KEY: Other
LOCATION: 1...18
OTHER INFORMATION: N-terminal acetylated and
OTHER INFORMATION: C-terminal amidated
US-09-465-719-215
Query Match 60.3%; Score 35; DB 4; Length 18;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 LOEWLKKLK 9
DB 10 LEWKKKKLK 18
RESULT 32
US-09-453-605-215
Sequence 215, Application US/09453605
Patent No. 6329341
GENERAL INFORMATION:
APPLICANT: Dasseux, Jean-Louis
APPLICANT: Sekul, Renate
APPLICANT: Butner, Klaus
APPLICANT: Cornut, Isabelle
APPLICANT: Metz, Gunther
TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
NUMBER OF SEQUENCES: 258
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,605
FILING DATE: 26-NO. 6329341-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/940,095
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 009196-0004-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6329341e
FEATURE:
NAME/KEY: Other
LOCATION: 1...18
OTHER INFORMATION: N-terminal acetylated and
C-terminal amidated
SEQUENCE DESCRIPTION: SEQ ID NO: 215:
US-09-453-605-215

Query Match 60.3%; Score 35; DB 4; Length 18;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKKLK 9
|:|:|:|:|
Db 10 LEEWKOKLK 18

RESULT 33
US-09-453-838-215
Sequence 215, Application US/09453838
Patent No. 6376464
GENERAL INFORMATION:
APPLICANT: Dasseux, Jean-Louis
APPLICANT: Sekul, Renate
APPLICANT: Butner, Klaus
APPLICANT: Cornut, Isabelle
APPLICANT: Metz, Gunther
APPLICANT: Dufourcq, Jean
TITLE OF INVENTION: APOLIPROTEIN A-I AGONISTS
TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
NUMBER OF SEQUENCES: 258
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,838
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/940,095

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 009196-0004-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6376464e
FEATURE:
NAME/KEY: Other
LOCATION: 1...18
OTHER INFORMATION: N-terminal acetylated and
C-terminal amidated
US-09-453-838-215

Query Match 60.3%; Score 35; DB 4; Length 18;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKKLK 9
|:|:|:|:|
Db 10 LEEWKOKLK 18

RESULT 34
US-08-437-841-14
Sequence 14, Application US/08437841
Patent No. 5563123
GENERAL INFORMATION:
APPLICANT: Innis, Michael
APPLICANT: Creasey, Abba
TITLE OF INVENTION: Chimeric Proteins
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.308
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/437,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/286,521
FILING DATE: 05-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Saveriede, Paul B.
REGISTRATION NUMBER: 36,914
REFERENCE/DOCKET NUMBER: 00990,001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-655-3542
TELEFAX: 510-601-2585
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-437-841-14

Query Match 60.3%; Score 35; DB 1; Length 20;
Best Local Similarity 60.0%; Pred. No. 59;
Matches 6; Conservative 2; Mismatches 0; Gaps 0;

Oy 1 LQEWLKKLKK 10
|:|:|:|:|
Db 5 LKKWLKMKFK 14

RESULT 35
US-08-286-521-14
Sequence 14, Application US/08286521
Patent No. 5589359

GENERAL INFORMATION:

APPLICANT: Innis, Michael

APPLICANT: Creasey, Adla

TITLE OF INVENTION: Chimeric Proteins

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESSES:

ADDRESS: Chiron Corporation

STREET: 4560 Horton St.

CITY: Emeryville

STATE: CA

COUNTRY: USA

ZIP: 94608

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30B

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/286,521

FILING DATE: 05-AUG-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Saveriede, Paul B.

REGISTRATION NUMBER: 36,914

REFERENCE/DOCKET NUMBER: 0990.001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 510-601-2585

TELEFAX: 510-655-3542

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-286-521-14

Query Match 60.3%; Score 35; DB 1; Length 20;
Best Local Similarity 60.0%; Pred. No. 59;
Matches 6; Conservative 2; Mismatches 0; Gaps 0;

Oy 1 LQEWLKKLKK 10
|:|:|:|:|
Db 5 LKKWLKMKFK 14

RESULT 36
US-08-436-175-14

Sequence 14, Application US/08436175

Patent No. 5696088

GENERAL INFORMATION:

APPLICANT: Innis, Michael

APPLICANT: Creasey, Adla

TITLE OF INVENTION: Chimeric Proteins

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESSES:

ADDRESS: Chiron Corporation

STREET: 4560 Horton St.

CITY: Emeryville

STATE: CA

COUNTRY: USA

ZIP: 94608

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30B

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/436,175

FILING DATE: 09-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/286,521

FILING DATE: 05-AUG-1994

ATTORNEY/AGENT INFORMATION:

NAME: Saveriede, Paul B.

REGISTRATION NUMBER: 36,914

REFERENCE/DOCKET NUMBER: 0990.001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 510-601-2585

TELEFAX: 510-655-3542

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-436-175-14

Query Match 60.3%; Score 35; DB 1; Length 20;
Best Local Similarity 60.0%; Pred. No. 59;
Matches 6; Conservative 2; Mismatches 0; Gaps 0;

Oy 1 LQEWLKKLKK 10
|:|:|:|:|
Db 5 LKKWLKMKFK 14

RESULT 37
US-08-435-149-10

Sequence 10, Application US/08435149

Patent No. 5866402

GENERAL INFORMATION:

APPLICANT: INNIS, MICHAEL A.

APPLICANT: ZAROR, ISABEL

APPLICANT: CREASEY, ADLA A.

TITLE OF INVENTION: CHIMERIC MCP AND DAF PROTEINS WITH CELL

TITLE OF INVENTION: SURFACE LOCALIZING DOMAIN

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESSES:

ADDRESS: CHIRON CORPORATION

STREET: INTELLECTUAL PROPERTY - R440, P.O. BOX 8097

CITY: EMERYVILLE

STATE: CALIFORNIA

COUNTRY: U.S.A.

ZIP: 94662-8097

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/435,149

FILING DATE: 05-MAY-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: SAVERIEDE, PAUL B.

REGISTRATION NUMBER: 36,914

REFERENCE/DOCKET NUMBER: 0989.001

TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2585
TELEFAX: (510) 655-3542
TELEX: N/A
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-435-149-10

Query Match 60.3%; Score 35; DB 2; Length 20;
Best Local Similarity 60.0%; Pred. No. 59;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKLKK 10
DB 5 LKKMLKMKK 14

RESULT 38
US-08-943-682-14
Sequence 14, Application US/08943682
Patent No. 6174721
GENERAL INFORMATION:
APPLICANT: Innis, Michael
APPLICANT: Creasey, Abia
TITLE OF INVENTION: Chimeric Proteins
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,682
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/438,184
FILING DATE: 09-MAY-1995
APPLICATION NUMBER: US 08/286,521
FILING DATE: 05-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Saveriede, Paul B.
REGISTRATION NUMBER: 36,914
REFERENCE/DOCKET NUMBER: 0990.001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2585
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-943-682-14

Query Match 60.3%; Score 35; DB 4; Length 20;
Best Local Similarity 60.0%; Pred. No. 59;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 LOEWLKLKK 10

DB 5 LKKMLKMKK 14

RESULT 39
PCT-US95-09464-14
Sequence 14, Application PC/TUS9509464
GENERAL INFORMATION:
APPLICANT: CHIRON CORPORATION
TITLE OF INVENTION: Chimeric Proteins
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09464
FILING DATE: 25-JULY-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saveriede, Paul B.
REGISTRATION NUMBER: 36,914
REFERENCE/DOCKET NUMBER: 0990.100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2585
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-09464-14

Query Match 60.3%; Score 35; DB 5; Length 20;
Best Local Similarity 60.0%; Pred. No. 59;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LOEWLKLKK 10
DB 5 LKKMLKMKK 14

RESULT 40
US-09-562-737-85
Sequence 85, Application US/09562737
Patent No. 6428967
GENERAL INFORMATION:
APPLICANT: Hertz, Joachim
APPLICANT: Gotthardt, Michael
TITLE OF INVENTION: LDL Receptor Signaling Pathways
FILE REFERENCE: UTSW0708
CURRENT APPLICATION NUMBER: US/09/562,737
CURRENT FILING DATE: 2000-05-01
NUMBER OF SEQ ID NOS: 132
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 85
LENGTH: 1024
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-85

Query Match 60.3%; Score 35; DB 4; Length 1024;
Best Local Similarity 63.6%; Pred. No. 1.6e+03;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LOEWLKKLKK 11
11:111111

DB 445 LQOMLEQLKKM 455

RESULT 41
US-09-238-303-9
; Sequence 9, Application US/09238303B
; Patent No. 6284253
; GENERAL INFORMATION:
; APPLICANT: Barr, Margaret C.
; TITLE OF INVENTION: No. 6284253el Feline Immunodeficiency Virus Nucleotide Sequence
; FILE REFERENCE: 18617.0059
; CURRENT APPLICATION NUMBER: US/09/238,303B
; CURRENT FILING DATE: 1999-01-28
; EARLIER APPLICATION NUMBER: US 60/072,927
; EARLIER FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 9
; LENGTH: 1150
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: protein encoded by the pol gene of a recombinant viral
; OTHER INFORMATION: clone constructed from the genomic DNA of a Pallas's cat feline
US-09-238-303-9

Query Match 60.3%; Score 35; DB 4; Length 1150;
Best Local Similarity 55.6%; Pred. No. 1.8e+03;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LOEWLKKL 9
1111111

DB 1029 LKEWIKR 1037

RESULT 42
US-08-193-521-17
; Sequence 17, Application US/08193521
; Patent No. 5470950
; GENERAL INFORMATION:
; APPLICANT: Maloy, W. Lee
; APPLICANT: Kari, U. Prasad
; APPLICANT: Williams, Jon I.
; TITLE OF INVENTION: Biologically Active Peptide
; TITLE OF INVENTION: Compositions and Uses Therefor
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DM4.V2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/193,521
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,960
; FILING DATE:

APPLICATION NUMBER: 07/760,054
FILING DATE: 13-SEP-1991
ATTORNEY/AGENT INFORMATION:
NAME: Olstein, Elliot M.
REGISTRATION NUMBER: 24,025
REFERENCE/DOCKET NUMBER: 421250-161
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: May be a C-terminal amide,
OTHER INFORMATION: and/or may be acetylated at
OTHER INFORMATION: N-terminus.
US-08-193-521-17

Query Match 58.6%; Score 34; DB 1; Length 11;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LOEWLKKLKK 11
11:111111

DB 1 LKLKLKKLKL 11

RESULT 43
US-08-434-120-111
; Sequence 111, Application US/08434120
; Patent No. 5635479
; GENERAL INFORMATION:
; APPLICANT: Baker, Margaret A.
; APPLICANT: Jacob, Leonard S.
; APPLICANT: Maloy, W. Lee
; TITLE OF INVENTION: Treatment of Gynecological
; TITLE OF INVENTION: Malignancies with
; TITLE OF INVENTION: Biologically Active Peptides
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DM4.V2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,120
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/297,950
; FILING DATE:
; APPLICATION NUMBER: US/08/226,108
; FILING DATE:
; APPLICATION NUMBER: US/07/937,462
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
NAME: Olstein, Elliot M.
REGISTRATION NUMBER: 24,025
REFERENCE/DOCKET NUMBER: 421250-194
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 111:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-434-120-111

Query Match 58.6%; Score 34; DB 1; Length 11;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKLKLM 11
1 LKLLKLKLL 11

RESULT 44
US-08-465-325-110

; Sequence 110, Application US/08465325
; Patent No. 5686563

; GENERAL INFORMATION:

; APPLICANT: Magalain Pharmaceuticals Inc.

; APPLICANT: 5110 Campus Drive

; APPLICANT: Plymouth Meeting, PA 19462

; TITLE OF INVENTION: Biologically Active Peptides Having

; TITLE OF INVENTION: N-Terminal Substitutions

; NUMBER OF SEQUENCES: 153

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Flanagan, Henderson, Farabow, Garrett &

; STREET: 1300 I. Street, N.W. Suite 700

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20005-3315

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/465,325

; FILING DATE: 05-JUN-1995

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/184,462

; FILING DATE: 18-JAN-94

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/891,201

; FILING DATE: 01-JUN-92

; ATTORNEY/AGENT INFORMATION:

; NAME: Fordis, Jean B

; REGISTRATION NUMBER: 32,984

; REFERENCE/DOCKET NUMBER: 05387, 0021-03000

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 408-4000

; TELEFAX: (202) 408-4000

; INFORMATION FOR SEQ ID NO: 110:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 11 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-465-325-110

Query Match 58.6%; Score 34; DB 1; Length 11;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKLKLM 11
1 LKLLKLKLL 11

RESULT 45
US-08-343-882-11

; Sequence 11, Application US/08343882
; Patent No. 5792831

; GENERAL INFORMATION:

; APPLICANT: Maloy, W. Lee

; TITLE OF INVENTION: Compositions of and Treatment

; TITLE OF INVENTION: With Biologically Active

; TITLE OF INVENTION: Peptides Having D-amino acid

; TITLE OF INVENTION: Residues

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carella, Byrne, Baln,

; ADDRESSEE: Giffillan, Cecchi, Stewart &

; STREET: 6 Becker Farm Road

; CITY: Roseland

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07068

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch diskette

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: DM4.V2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/343,882

; FILING DATE: 17-NOV-1994

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/133,740

; FILING DATE: 05-OCT-1993

; APPLICATION NUMBER: 07/874,685

; FILING DATE: 28-APR-1992

; APPLICATION NUMBER: 07/522,688

; FILING DATE: 14-MAY-1990

; APPLICATION NUMBER: 07/476,629

; FILING DATE: 08-FEB-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Olstein, Elliot M.

; REGISTRATION NUMBER: 24,025

; REFERENCE/DOCKET NUMBER: 421250-89

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-994-1700

; TELEFAX: 201-994-1744

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 11 amino acids

; TYPE: amino acid

; STRANDEDNESS: linear

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-343-882-11

Query Match 58.6%; Score 34; DB 1; Length 11;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKLKLM 11
1 LKLLKLKLL 11

RESULT 46
US-09-115-737-110

; Sequence 110, Application US/09115737
; Patent No. 6348445

; GENERAL INFORMATION:

APPLICANT: U. Prasad Kari
Taify J. Williams
Michael McLane
TITLE OF INVENTION: Biologically Active Peptides With Reduced
Toxicity in Animals and a Method for Preparing Same
NUMBER OF SEQUENCES: 156
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
Dunner, L.L.P.
STREET: 1300 I Street, N.W. Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/115,737
FILING DATE: 15-JUL-1998
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,330
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: 08/184,462
FILING DATE: 18-JAN-94
APPLICATION NUMBER: 07/891,201
FILING DATE: 01-JUN-92
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 05387.0021-06000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 110:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 110:
US-09-115-737-110

Query Match 58.6%; Score 34; DB 4; Length 11;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LOEWLKLKL 11
1 LKLLKLKL 11
DB

RESULT 47
US-08-465-325-133
Sequence 133, Application US/08465325
Patent No. 5686563
GENERAL INFORMATION:
APPLICANT: Magainin Pharmaceuticals Inc.
APPLICANT: 5110 Campus Drive
APPLICANT: Plymouth Meeting, PA 19462
TITLE OF INVENTION: Biologically Active Peptides Having
TITLE OF INVENTION: N-Terminal Substitutions
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
Dunner
STREET: 1300 I. Street, N.W. Suite 700
CITY: Washington
STATE: D.C.

COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,325
FILING DATE: 05-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/184,462
FILING DATE: 18-JAN-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/891,201
FILING DATE: 01-JUN-92
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 05387.0021-03000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 133:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-465-325-133

Query Match 58.6%; Score 34; DB 1; Length 12;
Best Local Similarity 63.6%; Pred. No. 52;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LOEWLKLKL 11
1 LKLLKLKL 11
DB

RESULT 48
US-09-115-737-133
Sequence 133, Application US/09115737
Patent No. 6348445
GENERAL INFORMATION:
APPLICANT: U. Prasad Kari
Taify J. Williams
Michael McLane
TITLE OF INVENTION: Biologically Active Peptides With Reduced
Toxicity in Animals and a Method for Preparing Same
NUMBER OF SEQUENCES: 156
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
Dunner, L.L.P.
STREET: 1300 I Street, N.W. Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/115,737
FILING DATE: 15-JUL-1998
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,330
FILING DATE: 05-JUN-1995

APPLICATION NUMBER: 08/184,462
FILING DATE: 18-JAN-94
APPLICATION NUMBER: 07/891,201
FILING DATE: 01-JUN-92
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 05387,0021-06000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 133:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 133:
US-09-115-737-133

Query Match 58.6%; Score 34; DB 4; Length 12;
Best Local Similarity 63.6%; Pred. No. 52;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LQEWLKLKLM 11
1 LKLKLKLKLM 11

RESULT 49
US-08-193-521-19
Sequence 19, Application US/08193521
Patent No. 5470850

GENERAL INFORMATION:
APPLICANT: Maloy, W. Lee
APPLICANT: Karl, U. Prasad
APPLICANT: Williams, Jon I.
TITLE OF INVENTION: Biologically Active Peptide
TITLE OF INVENTION: Compositions and Uses Therefor
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: New Jersey
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: DM4.V2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/193,521
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/870,960
FILING DATE:
APPLICATION NUMBER: 07/760,054
FILING DATE: 13-SEP-1991
ATTORNEY/AGENT INFORMATION:
NAME: Olstein, Elliot M.
REGISTRATION NUMBER: 24,025
REFERENCE/DOCKET NUMBER: 421250-161
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: May be a C-terminal amide,
OTHER INFORMATION: and/or may be acetylated at
OTHER INFORMATION: N-terminus.
US-08-193-521-19

Query Match 58.6%; Score 34; DB 1; Length 14;
Best Local Similarity 63.6%; Pred. No. 60;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LQEWLKLKLM 11
1 LKLKLKLKLM 11

RESULT 50
US-08-434-120-113

Sequence 113, Application US/08434120
Patent No. 5635479
GENERAL INFORMATION:
APPLICANT: Baker, Margaret A.
APPLICANT: Jacob, Leonard S.
APPLICANT: Maloy, W. Lee
TITLE OF INVENTION: Treatment of Gynecological
TITLE OF INVENTION: Malignancies with
NUMBER OF SEQUENCES: 117
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: New Jersey
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: DM4.V2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,120
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/297,950
FILING DATE:
APPLICATION NUMBER: US/08/226,108
FILING DATE:
APPLICATION NUMBER: US/07/937,462
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Olstein, Elliot M.
REGISTRATION NUMBER: 24,025
REFERENCE/DOCKET NUMBER: 421250-194
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-434-120-113

Query Match 58.6%; Score 34; DB 1; Length 14;
Best Local Similarity 63.6%; Pred. No. 60;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LOEWLKKKKM 11

1 : : | | | | | :
Db 1 LKKLLKKKKL 11

Search completed: July 11, 2003, 10:51:44
Job time : 33 secs